Teacher's study guide on the BIOLOGY of human populations

Asia

the unesco press

Teacher's study guide on the biology of human populations

Asia

Teacher's study guide on the biology of human populations

Asia

The Unesco Press
Paris 1975

Published by
The Unesco Press
7 Place de Fontenoy, 75700 Paris
Printed by Imprimerie des
Presses Universitaires de France, Vendôme

ISBN 92-3-101279-7
French edition 92-3-201279-0

Preface

The main objective of Unesco's pre-university science and technology education programme is to organize and promote all activities leading to the improvement of the teaching of science disciplines at pre-university level. The Organization's aim in this field, particularly in developing countries, has been to introduce improved teaching programmes for each particular discipline on a regional basis.

As regards biology, the African continent was chosen for the implementation, between 1967 and 1972, of a pilot project for the improvement of education in this discipline. Important results have been obtained thanks to the co-operation of the national study groups set up in African countries, training courses and seminars lasting from a few weeks to several months and, finally, the co-operation of experienced African and foreign specialists and educators. Among these results should be mentioned: publication of several teacher's guides with colour slides adapted to African needs; advice on methods of improving biology curricula; provision of plans for inexpensive laboratory experiments; encouragement to countries to adapt the pilot project materials and methodology more closely to their national requirements (several countries have already commenced such work).

Moreover, the dissemination of documents on biology teaching by Unesco with the help of regional and governmental organizations, has certainly led to greater awareness of national efforts in the field of educational activities. This has also been of benefit to Asia and Latin America, whose countries can profit from the experience acquired and the methods employed.

In order to support this national effort in educational renewal, which covers not only the improvement of the content and methods of teaching programmes, but also of teacher training, and above all the adaptation of education to its social and cultural context, Unesco undertook in 1971–72 to prepare a teacher's study guide on several aspects of the biology of human populations.

Human biology forms a part of the biology curriculum in secondary schools at all levels. It is a complex subject and sometimes presents difficulties for the teacher who often has little documentation available related to local conditions, while at the

same time it raises a number of crucial questions on the behaviour of individuals and the evolution of society. One of the courses organized within the framework of the Pilot Project for Biology Teaching in Africa was devoted to the preparation of teaching material on some aspects of human biology (nutrition, reproduction and diseases), intended for hygiene and nutrition education, emphasis being placed on the physiological aspects. It seemed important to extend this initial undertaking to other continents, at the same time giving it a more ecological dimension; that is to say, dealing with the biology of human populations rather than of man alone. This initiative, moreover, came at a time of world-wide concern with the problems of environment and of relations between man and his milieu which were being stated everywhere in critical and sometimes dramatic terms. It was therefore natural to try to introduce this concern in secondary education programmes, by means of clear and objective information centred upon the larger theme of human ecology. It fell within Unesco's field of competence to undertake this initiative, in close co-operation with the United Nations Fund for Population Activities (UNFPA). Demographic growth and its consequences and human reproduction must, in fact, be seen in their biological and ecological context in order to be better understood at the sociological level. The assistance given by the UNFPA has been valuable and, indeed, vital in the completion of the present work.

The preparation of this book comprised three main stages. In 1972, three professors from American universities drafted the basic text containing the fundamental data on human genetics and the evolution of Man, the physical and biological environment of human populations, human reproduction and sexuality, as well as population dynamics, the relationships between populations and their environment (housing, degradation of the environment, pollution, etc.). At the end of 1972 and the beginning of 1973, three regional seminars were held in Montevideo, Nairobi and Bangkok, where, in the presence of the authors of the basic text and in collaboration with regional specialists and educators, the form and content of a guide more suited to the region concerned were decided. Finally, in 1973, the writing of the African, Asian and Latin American versions of the book was completed under the guidance of a regional co-ordinator in each case and, in the first half of 1974, the English version of these three texts was finalized at Unesco Headquarters. French and Spanish translations were completed during the second half of 1974 and in 1975.

It should be stressed that in these three versions—African, Asian and Latin American—this Teacher's Study Guide on the Biology of Human Populations satisfies a definite need, expressed during the three regional seminars, and appears at an opportune moment. It is the result of collective work, and therefore necessarily reflects, with undoubted objectivity, the nuances of different modes of thought. For this reason, it is designed for use only by an experienced teacher, who will make best use of it by adapting the information he obtains from it to the needs of his country, society and students. The directions which he will find in the Introduction immediately following this preface will help him in this respect. The Guide, moreover, makes no claims to be an exhaustive work, nor is the effort of adaptation complete. It would not be possible to cover all aspects of this vast subject and to take into account all the particular

situations encountered throughout a continent. For this reason, selected bibliographical references have been placed at the end of each Part of the book, in order to guide further investigation on the part of the teacher. The latter should, moreover, consult local documentation sources (universities, ministries of technology, research centres) and regional and international organizations (publications issued by Unesco, FAO, WHO, UNFPA and their regional offices; publications issued by the United Nations, the Population Council, etc.). The numerous illustrations in this Guide should also prove very useful, especially in those countries where educational resources are still limited at secondary level; their origin is indicated in each case, which means they are an important source guide for documentation.

The effort to adapt the Guide has received particular attention and, given the enormity of the task, may be considered satisfactory. It is the sincere hope of Unesco and UNFPA that this effort will be continued in each country concerned, with a view to producing similar guides at national and sub-regional level. This, moreover, constitutes the second stage of the project: the organization of regional seminars during which the work will be still better adapted with the collaboration of the experts in the region or sub-region. Assistance—although it may be modest—will continue to be provided to those national study groups or commissions for teaching reform who wish to undertake such a task.

The designations employed and the presentation of the material in this work do not imply the expression of any opinion whatsoever on the part of the Unesco Secretariat concerning the legal status of any country or territory, or of its authorities, or concerning the delimitation of its frontiers.

The opinions expressed in the following pages, moreover, are those of the authors and do not necessarily reflect those of Unesco.

Acknowledgements

Unesco is grateful to Professors LeVon Balzer and I. L. Slesnick (Biology Department, Western Washington State College, Bellingham, Washington) and Professor Kenneth Linton (Clarion State College, Clarion, Pennsylvania), who were the authors of a basic text; to Professor F. C. Vohra (School of Biological Sciences, University of Malaya, Kuala Lumpur), who contributed chapters and, as regional co-ordinator, was also responsible for the Asian version; to Dr W. D. Ponniah (Research Officer, RECSAM, Penang) and Mr G. Guru (National Institute of Education, National Council of Education, New Delhi), who also contributed some chapters; and to Messrs J. F. Banaagh (Philippines), K. Ohkura (Japan), M. R. Puttipongse Varavudhi (Thailand) and M. Soraya (Iran), who participated in the Bangkok seminar, during which the general layout of the book was prepared. The final editorial work was completed by Professor A. Sasson (Department of Plant Sciences, Faculty of Science of Rabat, Morocco).

Unesco is also grateful to the authors and publishers of works from which various illustrations and numerous data were taken.

Table of contents

Introduction
The use of this book 13
Ecology: concepts and principles 16
Biosphere 20

Part I *Evolution of human populations*
1 The evolution of man 39
2 Genetics background 53
3 Mechanisms of evolution 67
Further reading 78

Part II *Environment of human populations*
4 Energy 81
5 Atmosphere 117
6 Water 139
7 Soils 173
8 The biotic environment: flora and fauna 213
9 The oceans 231
10 Food and nutrition 245
11 Diseases and environment 271
12 Mineral resources 285
Further reading 302

Part III *Dynamics of human populations*
Preamble 307
13 Demographic parameters 309
14 Population dynamics: growth and density trends 327
15 Environment and population patterns 341
Further reading 352

Part IV *Reproduction in human populations*
Preamble 355
16 Human reproduction 357
17 Sexual behaviour 381
Further reading 390

Part V *Present and future evolution or Design for survival*
18 Ecosystem management and control of environmental quality 393
19 Regulation of fertility 413
20 Does man have a future? 433
Further reading 437

The use of this book

This book has been written to help teachers to cope with all the new aspects related to man and to his action on the environment. It is a source of information on biological subjects not usually treated in student texts because, only in the last years, have they become related to important issues of our times. A minimum of materials already in conventional use in secondary teaching had to be included, in order to give unity to the presentation.

This is not a student text. It is a teacher's study guide. Thus, its direct consultation by secondary students is clearly *not* recommended. In order to make it more relevant to teachers, the authors have adhered to the following guide-lines: to treat every subject freely and impartially, as fits a scientific discussion; to offer factual and conceptual information without indulging in propagating unsubstantiated opinion.

Good techniques for teaching sciences are today well divulged. They certainly differ greatly from the traditional teaching to which many of us have been exposed: namely, lecturing by the teacher and note-taking by the student. We emphatically advise teachers not to make the substance of this book the object of lecturing, under the excuse, for instance, that its contents are important and have not been included in a student's book. The following suggestions for its proper use are therefore made.

The use of the teacher's study guide

First, we hope that the guide will help teachers to widen their own attitudes and academic experience so that they can conduct their courses more fruitfully by modern methods. Furthermore, they will be free to make written summaries based on it, adapted to the level of their classes, and mimeograph or otherwise reproduce them for the direct use of the students. In doing this, it is important that the material be presented in a more stimulating style than the one used here, so that giving information becomes an aim subordinated to creating interest, raising problems and giving hints for their discussion. Guide-questions for the student to answer in writing will help focusing his attention to

the crucial aspects of the subject and will make both his work and the discussion which must follow more productive.

The way for a meaningful teaching of biology at the secondary level is to place under the student's scrutiny: questions about himself, as a living being; biological problems of practical importance for his country or community, and issues of great interest for mankind in general. The subjects presented here relate mainly to such categories. It is not sufficient, however, to choose pertinent materials; it is essential that they be thoroughly relevant to the surroundings of the student and to his past experience. The best way to do this is to adopt the problem method in either of its variants, i.e. the discussion and the project techniques.

The problem method

Here are some of the abilities the problem method helps students to develop:

1. Drawing information from books, encyclopaedias and other sources, when needed, for the development of the subject he is studying.
2. Making summaries and drawing conclusions from readings and discussions, discarding trivial material and concentrating on parts which are important for his purposes.
3. Making use of basic equipment such as microscopes, centrifuges and glassware and manipulating chemicals safely and competently.
4. Making use of basic mathematical techniques when necessary.
5. Planning to conduct interviews and inquiries to obtain information.
6. Making a critical judgement of statements and reasonings from books and people.
7. Explaining ideas clearly, in writing or orally.
8. Planning and performing observations and experiments.

Among the mental attitudes which the problem method develops are the following:

1. To look for factors influencing a situation as a means of understanding it, for cause and effect relationships through valid criteria.
2. To avoid hasty judgement and unwarranted generalizations, and to base opinions on proved facts as far as possible.
3. To look for biases in one's own opinions and reasonings and examine other people's ideas openmindedly;

therefore to be ready to revise opinions as soon as new evidence is presented.

4. To judge the efficiency and precision of the methods and techniques used by oneself and others in collecting evidence.
5. To preserve his curiosity and interest with respect to all important aspects of science and its application.
6. To use scientific knowledge and the scientific way of thinking in current personal, communal and professional life.

These attitudes and principles show that the important thing is that the student improves the efficiency of his thinking during his training. The task of the teacher is mainly to excite the minds of the students through worth-while demonstrations, laboratory work, excursions, readings and discussions.

The whole process becomes more natural and efficient when students work in groups. Preparatory work, such as reading and summarizing, collecting information in books and interviewing people will be better done by small teams of students; the discussion of data gathered, answers to guide-questions, and reasonings and conclusions produced is more profitable if all in the class participate.

Some subjects, like reproduction in humans, are appropriate to be worked on through the discussion technique, variant of the problem method, while others, like pollution, will merit the use of the project technique.

The discussion technique

The efficiency of this teaching procedure depends specially on the adequacy of the written sources made available to the student and the skill of the teacher in conducting discussions. Basically, the ensuing steps may be followed:

1. The teacher presents to the student, in no more than ten minutes, the nature of the problems to be studied, being careful to emphasize the important points, without stating their solutions. This raises interest and helps the student perceive the aims of his work.

2. During one or more sessions the student reads and discusses in small groups some appropriate material and answers the guiding questions. This phase can be also worked out individually, in class or at home. The teacher makes sure that each team (or each student) writes down its (or his) answers to each guiding question.

3. To start the general discussion the teacher asks a student his answer to the first guiding question and the other students are encouraged to comment on it. Question after question is thus discussed until there is agreement in the class about the answer to each. The teacher should not lose the opportunity for extending the discussion to cover subjects not included in the guiding questions.

The project technique

Typical of this technique are the following aspects:

1. The activity must be directed toward a well-defined goal, and the scope of the problem must be wide; for example, 'to what extent is BCG vaccination efficient in the prevention of tuberculosis?' It must be divisible into significant subproblems ('what are the causes of tuberculosis?') and tasks ('find out how often BCG vaccination is used in our community').
2. The aim of the project should not vary, but the means may change as the actual planning is frequently revised.
3. The fundamental part of the project is the practical work (observations, experiments, interviews, data-gathering) and this must culminate in some concrete manner such as a final report, an exhibit, a demonstration or a seminar.
4. The richness of a project is measured by its interaction not only amongst the students but also with the community.

Ecology: concepts and principles

Ecology

Today, the word 'ecology' is much used in many situations and on many occasions. However, the term is often applied rather indiscriminately and sometimes even considered synonymous with 'environmental science'. But there is a difference. Ecology, as originally defined by the German biologist Ernst Haeckel, is concerned not only with the environment, but also with the organisms in that environment and 'all the complex interrelations' referred to by Darwin as 'the conditions involved in the struggle for existence'.

The subject-matter currently embodied in the discipline of ecology is not new. Its viewpoint, in so far as its implications in the understanding of global problems together with its conception of the earth as a set of interlocking and interdependent systems; and its emphasis on man as an integral part of these systems, is now accepted and of great significance.

Consequently, ecology is no longer merely a subsection of biology but also an extension of social sciences and the humanities. Man's relationships with both animate and inanimate nature reflect on and interact with his socio-economic structures. Crucial questions of economic development, industrialization and population control, etc., are all bound up with the ecological problem. In brief, the 'holistic concept' is gaining foothold, developing integrative approaches to human problems and creating inter-

national actions and proposals for the solution of such problems. It is, therefore, imperative that the basic concepts of ecology are properly understood and acted upon by as many people as possible.

Living systems

The universe consists of systems, each higher level of system the components of which are of units of the lower level, e.g. atoms are composed of elementary particles; molecules, of atoms; and organelles of molecules, and so on. A system can be defined as a set of regularly interacting and interdependent component forming a unified whole. The word set implies that the units have some common properties. The definition implies homeostasis and feedback. 'Every system', according to Ramon Margalef, 'is a set of different elements or units or compartments any one of which can exist in many different states, such that the selection of the state is influenced by the states of the other components of the unit. Elements linked by reciprocal influences constitute a feedback loop.' A system is either closed or open. The latter has boundaries which are at least partially permeable to matter-energy or information transmissions of some sort. The boundaries of a closed system, on the other hand, are impermeable. However, there is no system which is either completely closed or completely open. It is either relatively open or relatively closed. Each system has some intrinsic matter-energy. With the passage of time the energy is slowly used up and the matter gradually becomes disorganized.

If the system is made up of living components, it is called a living system. Such systems have the following characteristics: they are open systems; they require inputs of food or fuels to restore spent energy and components; they have a certain degree of complexity; they are largely composed of protoplasm; they contain genetic material; they have subsystems one of which is the crucial and causes others to interact; they have symbiotic, parasitic or similar relationships with other living or non-living systems; they exist only in a certain environment. Any drastic, severe, adverse change in any factor of the environment may prove destructive to the living system.

Thus to define modern ecology, in the words of E. P. Odum, is to consider the word in terms of the concept of level of organization visualized as a sort of 'biological spectrum' as shown in Fig. 1.

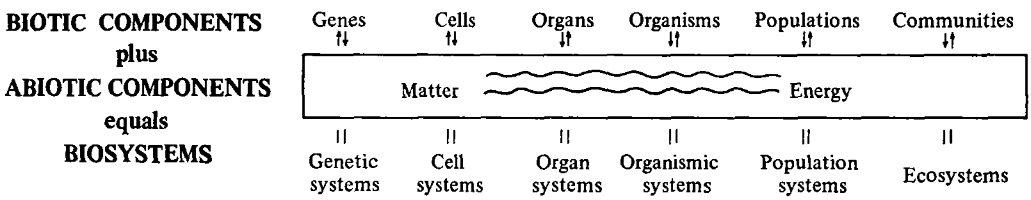


Fig. 1
Levels of organization spectrum. Ecology focuses on the right-hand portion of the spectrum, that is the levels of organization from organisms to ecosystems

Source: E. P. Odum. *Fundamentals of Ecology*. 3rd ed., Philadelphia, W. B. Saunders, p. 5.

Accordingly, the study of all system levels, from organisms and beyond on the right-hand side of the spectrum, would form the subject-matter of ecology. Regardless of precise definition, the substance of ecology is seen in a myriad of abiotic and biotic mechanisms and interrelations involved in moving energy and nutrients, regulating population and community structure and dynamics.

Living systems exist in various levels of organization, the lowest being that of cells as shown in Fig. 1. Cells are composed of atoms, molecules and organelles; organs are composed of cells aggregated into tissues and organisms of organs. This complexity extends to the levels of population, community, ecosystem and ultimately to one biosphere or one living earth also referred to as the ecosphere.

Population and community

A population is a group of individuals of the same species occurring in a distinct area. A species is a group of individuals of the same kind and capable of interbreeding. Similarly a community is an assemblage of all populations of animals (including man) and plants occupying a given area; and therefore interacting with one another. Maintenance of a community is dependent upon the flow of energy through functional strata of populations. Thus in each community, most populations can be assigned to one of the following roles, each known as a niche.

1. Producers—plants species which, through photosynthesis, convert solar energy into chemical energy contained in plant tissue.
2. Consumers of first order—herbivores, i.e. animals which ingest plants. They may be microbes (e.g. *Amoeba*) which feed on green plant cells or clams (Mollusca) which filter microscopic plants from the surrounding medium; they may be insects, mammals of similar forms feeding on some foliage or weeds.
3. Consumers of second order—carnivores, i.e. animals which feed on herbivores.
4. Consumers of third order—carnivores which feed on other carnivores. However, some animals including man are carnivores and can be classed as '2' '3' or '4' depending upon the situation.
5. Parasites—organisms which obtain their food directly from the bodies of other organisms in which or on which they live. The food, in this case, is predigested.
6. Scavengers—animals which feed on faeces and or on dead and decaying bodies of others. Unlike parasites, they digest their food in their gut.
7. Decomposers—most fungi, bacteria and microbes,

which decompose dead organisms and organic debris to release basic chemical substances into the environment eventually to be absorbed by living plants.

Ecosystems

A combination of a community and its non-living environment operating as a functional system is termed an ecosystem. It is the basic structural and functional unit of nature on the earth's surface. V. N. Sukachev calls it 'biogenocenose' and defines it as a genetically, geographically and trophically linked local combination of vegetation, animals, soils, relief, climate and hydrography. In other words 'they are parts of the land or the water surface, homogenous in respect to topographic, microclimatic, botanical, zoological, pedological, hydrological and geochemical conditions'. These are of many sizes and kinds. The concept that all organisms are interdependent and interrelated with one another and with their environment applies to all ecosystems from that of, say, a small pond to that of the world at large.

The 'eco' part of the word ecosystem implies environment which comprises the sum of all the external factors, processes and conditions that affect a living system. These can be either biotic (other living systems) or abiotic (non-living factors e.g. light, temperature, pressure, moisture, soil, nutrients, gases and fire, etc.). The two components, abiotic and biotic, interact continuously and influence each other's properties and both are necessary for maintenance of life on the earth.

The 'system' part, as indicated earlier, 'represents an interacting and interdependent complex'. Of all levels of organization, the ecosystem is the only self-sustaining one, requiring solar energy as the only external resource, other levels being dependent upon other life and environment for existence, in fact, on the ecosystem of which they are the part.

Limiting factors

In any environment, living things are distinguished by their growth, reproduction and mobility apart from other characteristics. Every living species tends to multiply and spread to new and suitable surroundings, repeating the process after establishment. Growth in population size continues usually until it is checked by some external factor. Such a factor, whatever its nature, is called a limiting factor. According to Odum, any condition that approaches or exceeds the limits of tolerance of an organism is said to be a limiting condition or a limiting factor. Limiting factors can be either physical (e.g. climate, presence or absence of water and nutrients, etc.) or biological (e.g. competition, predation, parasitism and disease, etc.).

Odum emphasizes no sharp lines or breaks in the spectrum (Fig. 1), not even between the organism and the population. This follows the principles of interdependence, interrelations and survival. He further states that the individual organism

cannot survive for long without its population any more than the organ would be able to exist for long as a self-perpetuating unit without its organism. Similarly the community cannot live without the cycling of materials and the flow of energy in the ecosystem. The largest and most nearly self-sufficient living system is called the biosphere.

Biosphere

According to Dasmann the biosphere can be considered as a term defining the sum total of all the ecosystems of the earth. Each ecosystem is physically and biologically different, with various organisms well adapted to the environment in which they occur. Despite this great diversity there is an underlying unity. Life, whereever it may exist, is organized on the same basic principles and various ecosystems (oceans, lakes, forests, grasslands and deserts, etc.) all operate in the same way, i.e. energy fixed by plants flows through them; nutrients are deposited in the tissues of organisms, cycled from one group to another; and finally released by decomposition to soil and water for recycling. Living organisms under natural conditions never exist in complete isolation. A change in any one component evokes changes in all the others. Thus, the orderliness of life is maintained through functional interdependence and interrelations at all levels. This orderliness, in turn, leads to mechanisms of checks and balances, actions and reactions which tend to maintain a steady state system. This tendency toward constancy or self-regulation is called homeostasis and involves the principle of controlled 'feedback' in which deviation of the system from the normal sets off counteractions that tend to bring the system back to its original state.

The concept of the biosphere accepted today was developed by the Russian geochemist V. I. Vernadsky, though its idea as collective totality of living creatures on the earth was first conceived by the French naturalist, Lamarck. According to Vernadsky, in the definition of biosphere, it is not sufficient to include only a space where life exists. He emphasizes the following three main components, basic to the concept of the biosphere: living matter—totality of organisms determined quantitatively by the biomass; biogenic matter—organo-mineral and organic products created by living matter, e.g. coal, bitumen, combustible gases, peat, soil humus and litter, etc.; biocosmic matter—mineral materials formed by the association of living organisms and non-living nature, e.g. the gas composition of the lower layers of the atmosphere, sedimentary rocks, clay minerals and water, etc.

Evelyn Hutchinson refers to the irregular vertical extent of the biosphere and traces its chronology (as represented in the fossil record), which are illustrated in Figs. 2 and 3. It has three realms—atmosphere, hydrosphere and lithosphere, which are arranged in order of their relative densities. The lithosphere is the basal solid rock

forming an irregular and uneven layer of large depressions and elevations. Our continents and islands are all part of this system. The hydrosphere represents the water masses forming a system next in decreasing density and fills the depressions between the continental plateaux, leaving uncovered the upper parts (i.e. the exposed

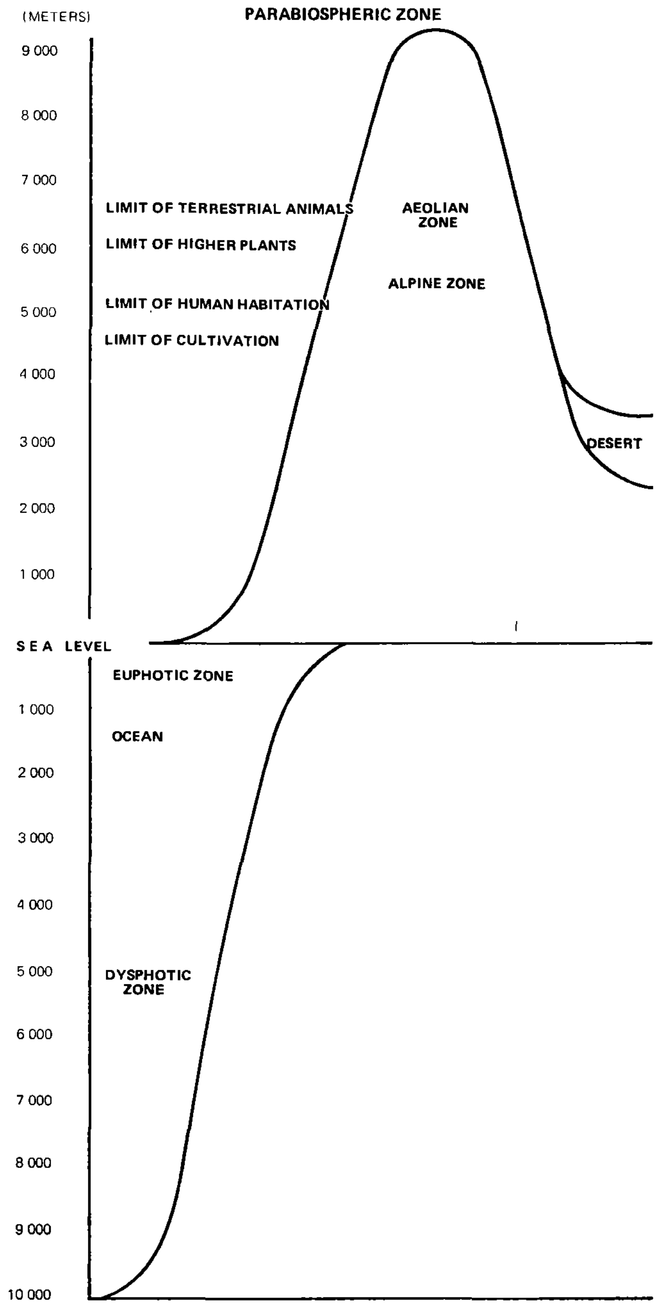


Fig. 2
Vertical extent of the biosphere illustrated schematically
Source: G. E. Hutchinson. 'The Biosphere', *Scientific American* (New York), September 1970, p. 6.

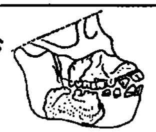
YEARS BEFORE PRESENT	EVENT	GEOLOGICAL FORMATION	FOSSIL
0	OLDEST HOMINID	SIWALIK HILLS (INDIA)	*RAMAPITHECUS* 
	OLDEST LAND PLANT	LUDLOVIAN SERIES, UPPER SILURIAN (BRITAIN)	*COOKSONIA* 
	OLDEST METAZOAN ANIMAL	EDIACARA HILLS (AUSTRALIA)	*SPRIGGINA* 
1 BILLION			
	OLDEST EUCARYOTIC CELLS	UPPER BECK SPRING DOLOMITE (CALIFORNIA)	*UNNAMED* 
	FORMATION OF OXIDIZING ATMOSPHERE		
2 BILLION	OLDEST PHOTOSYNTHETIC AND NITROGEN-FIXING ORGANISM	GUNFLINT FORMATION (ONTARIO)	*GUNFLINTIA* 
3 BILLION	OLDEST KNOWN ORGANISM	FIGTREE FORMATION (SOUTH AFRICA)	*EOBACTERIUM* 
	FIRST ROCKS IN EARTH'S CRUST : FORMATION OF OCEAN		
4 BILLION	DIFFERENTIATION OF EARTH'S CRUST, MANTLE AND CORE ; CRUST MELTED BY RADIOACTIVE HEATING		
4.5 BILLION	FORMATION OF EARTH		

Fig. 3
Rough chronology of the biosphere as presented in the fossil record. N.B. In the diagram, one billion equals 1,000 million

Source: G. E. Hutchinson. 'The Biosphere', *Scientific American* (New York), September 1970, p. 4.

land mass). The atmosphere is the gaseous envelope above the hydrosphere and touches the lithosphere only on the continents and the islands. Fig. 4 illustrates the simplified structural aspect of the biosphere. Life is made possible on the biosphere by virtue of three important conditions: existence of liquid water in substantial quantities, supply of energy from the sun, and presence of interfaces between the three realms of the biosphere.

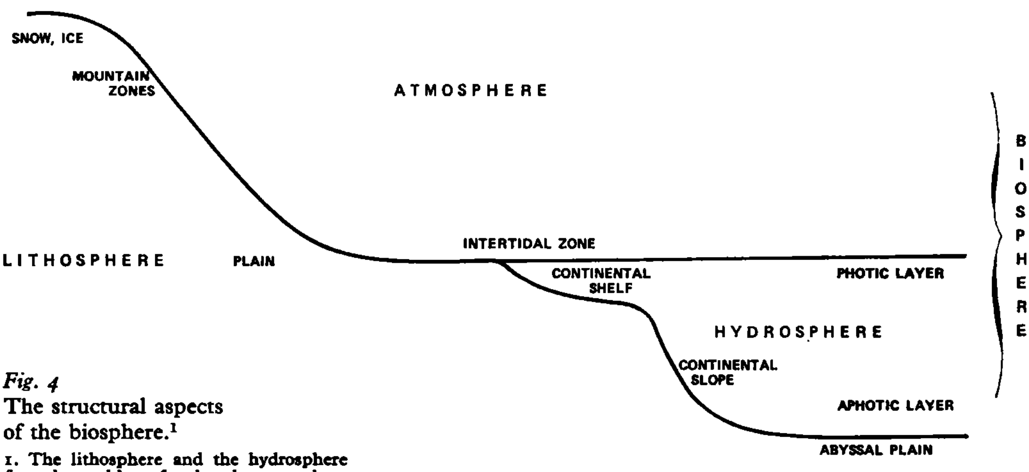


Fig. 4
The structural aspects of the biosphere.[1]

1. The lithosphere and the hydrosphere form the earth's surface but the atmosphere has no distinct limit.

The present state of the biosphere is the result of a long-term evolution. Since its estimated inception some 6×10^9 years ago, the earth's surface has been subject to the influence of energy from space. This has effected most complicated changes resulting in diversification of both abiotic and biotic components of the biosphere. An estimated number of about 3 million species of animals, plants and micro-organisms is said to have come into existence. All of them have acted as important biogeochemical agents in bringing about large-scale transformations in the earth's crust, which in turn have triggered changes of similar dimensions in the biotic world. Thus we have today, on the surface of the earth, structures such as the oceans, continents, lakes and glaciers of high-altitude regions and of the Arctic, etc., with their own specially adapted biota.

Role of living matter

Among others, the role of living matter in atmospheric, hydrological and soil processes includes: absorption and storage of mineral compounds from their geological cycles during nutrition and thus preventing their drainage from the land to the sea. Examples include enormous quantities of hydrocarbon compounds of the fossil fuels and the carbonate sediments; exchange of oxygen, carbon dioxide, ammonia and water vapour, etc., with the environment during their metabolic activity; weathering, migration and sedimentation of substances by their existing activities; reduction of

minerals and organic compounds mainly through micro-organisms; concentration and secretion of calcium in the form of carbonates, phosphates and oxalates, etc.; increasing biogenic concentration in soils.

In each case, the impact of the biota on various geochemical, atmospheric, hydrological and soil processes makes an impact throughout the cycles of matter and energy. Such cycles, in general, are open and unidirectional. During the course of nutrition and mineralization certain quantities of elements escape from their respective cycles and are locked away for various periods in oceanic sedimentary deposits.

Organisms differ in the type of energy and matter needed to sustain them in the biosphere. Accordingly, two classes of organisms: autotrophs and heterotrophs, are recognized. Autotrophs synthesize their own food from simplest forms of matter (carbon dioxide and minerals) using solar energy. Heterotrophs cannot make use of sunlight directly and instead depend upon chemical energy in the form of reduced carbon compounds or some specific amino acids of organic origin. Hence, the energy and matter inputs to organisms may be either inorganic (more important for autotrophs) or organic (more important for heterotrophs).

Matter and energy

Matter implies atoms and molecules, both unchanged and changed as ions, occurring singly or as compounds, or as aggregation of compounds. All matter possesses mass and occupies space. It exists as solid, liquid or gas. It exhibits gravitation. Energy, on the other hand, is the capacity to do work. Various forms of energy are recognized but those of greatest importance to living systems are mechanical, chemical, radiant and heat energy. Energy can be neither created nor destroyed in the universe. However, it may be converted from one form to another with some measurable loss, as heat, at each transformation. Matter may have kinetic energy, when it is moving and exerts a force on other matter; potential energy, attributed to it by virtue of its position in a gravitational field; or rest-mass energy which would be released if mass was to be converted into energy. Both mass and energy are interchangeable. Albert Einstein gave the exact relationship of matter to energy in his equation $E=MC^2$, where E is the quantity of energy, M is the quantity of mass and C is the speed of light. The speed of light, 300,000 km per second, is an enormous quantity.

Accordingly, a very minute quantity of mass, on conversion, will produce a tremendously large amount of energy. For example, the mass which is converted to energy in the fission of one pound of uranium is equivalent to the energy released by the burning of over 1,500 tons of coal.

Steady state

All living systems require an input of both matter and energy in order to synthesize and reproduce biological materials, and to maintain their own necessary internal energy levels. Matter and energy are also emitted from the system to achieve a balance or a steady state (Fig. 5). A steady state results when opposing variables in a system are in balance, i.e. when rate of input of energy and matter equals rate of output of energy and matter, when the storage of energy and matter remains unaltered. A system in a steady state is said to be in equilibrium with regard to such variables. The equilibrium may be static or dynamic. A living system is of latter type. Since it is in a continuous flux, a moderate alteration in one variable may produce greater or lower changes in other related ones. Such changes may or may not be reversible.

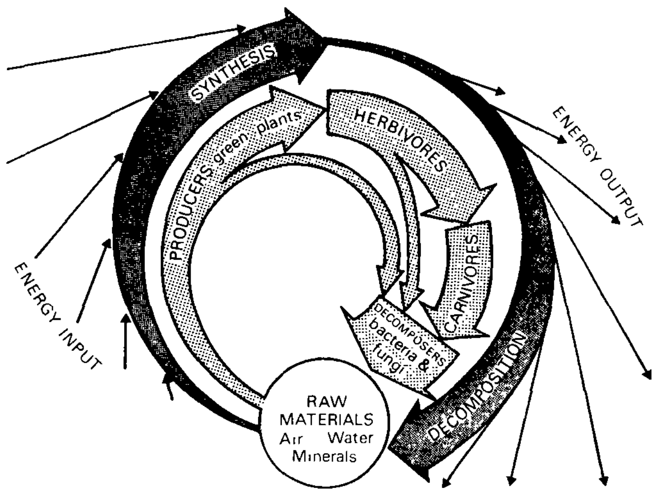


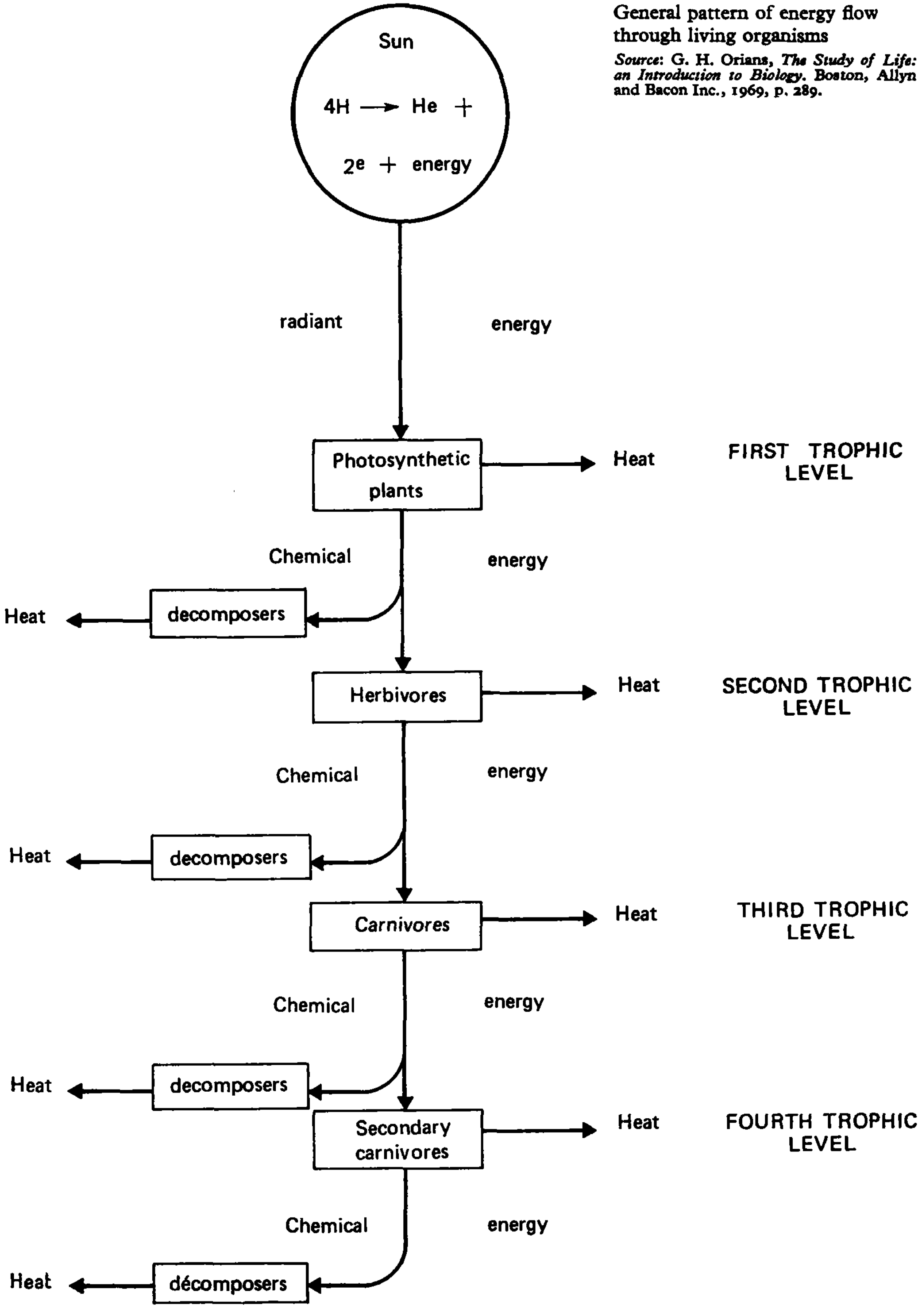
Fig. 5
Balanced cycle of synthesis and decomposition, showing the recycling of matter powered by the one way flow of energy
Source: R. B. Platt; George K. Reid. *Bioscience*, New York, Reinhold Publishing Corporation, 1967, p. 177, Fig. 8–5.

Pathway of energy

The biosphere can be visualized as a single great energy system which receives essentially the solar energy forming about 99.98 per cent of the total input. This amounts to about 1.3×10^{23} kc per year but only a very minute fraction of this energy (about four-hundredths of 1 per cent) is utilized in photosynthesis for a sustenance of the biosphere. The remainder is lost either through reflection back to space or absorption by the atmosphere.

With the fixation of solar radiation in plants, the energy flow begins and follows a general pattern, as shown in Fig. 6, through the rest of the living system. This movement is unidirectional with the amount of energy diminishing at each transformation until all is lost as waste heat. Figs. 7 and 8 show the general distribution and utilization of solar energy. Almost five-sixths of the solar energy trapped by plants is available to animals and other consumers. Man takes about a third of his share in food and the rest in mechanical, thermal and electrical energy used in his work.

Fig. 6
General pattern of energy flow through living organisms
Source: G. H. Orians, *The Study of Life: an Introduction to Biology*. Boston, Allyn and Bacon Inc., 1969, p. 289.



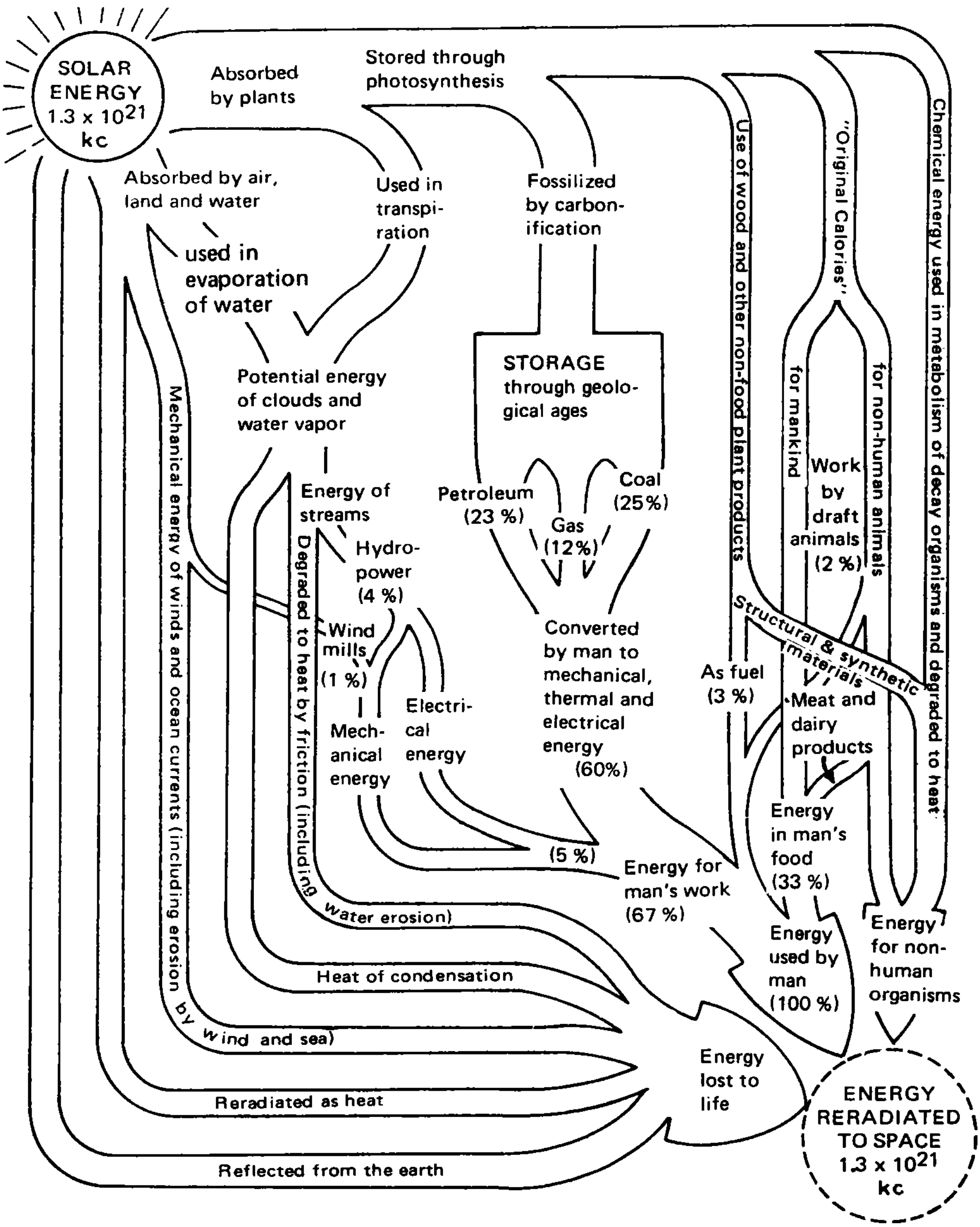


Fig. 7
Generalized distribution of solar energy in the biosphere

Source: L. J. Milne; M. Milne. *The Biotic World and Man*. Englewood Cliffs, New Jersey, Prentice Hall, 1965, p. 285.

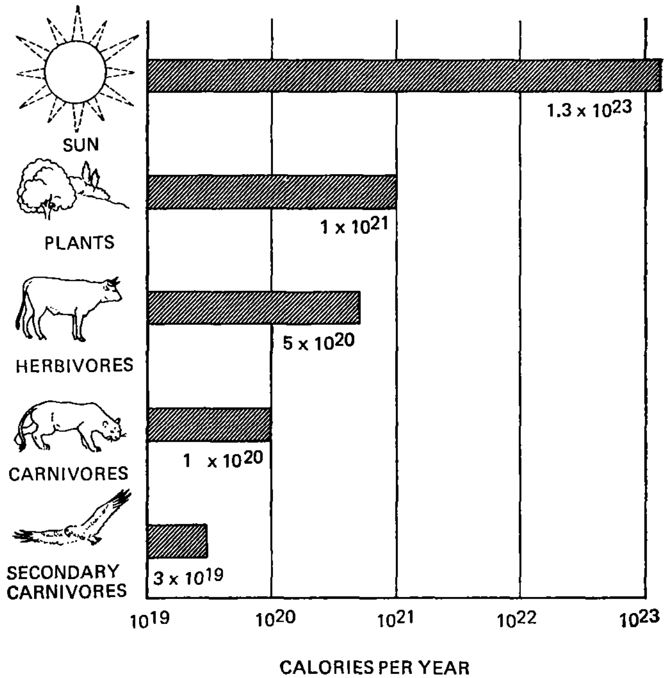


Fig. 8
Utilization of solar energy indicates decrease with each step along the food chain. The bars are shown on a logarithmic scale
Source: L. C. Cole. 'The Ecosphere', *Scientific American* (New York), April 1968.

Food chains and food webs

Most of the interactions between species involve food; competing for food, eating and avoidance of being eaten are the most common ways in which animals affect one another. The green plants, however, synthesize their own food. As such, the following illustration represents a simple feeding relationship in a biotic community:

$$\begin{array}{ccccccccccc} \text{Loss of heat energy} & & & & \text{Loss of heat energy} & & \text{Loss of heat energy} & & \text{Loss of heat energy} & & \text{Loss of heat energy} \\ \uparrow & & & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\ \text{Grass} & \longrightarrow & \text{grasshopper} & & \longrightarrow & \text{frog} & \longrightarrow & \text{snake} & \longrightarrow & \text{hawk} & \longrightarrow \text{bacteria of decay.} \end{array}$$

Such a pathway of food consumption is called a 'food chain'. In general there are three types of food chains: those starting from a plant base and going from smaller to larger animals (predators chains); those associated with parasites and consisting of smaller animals preying on large ones; and those forming saprophytic chains where dead matter is converted into micro-organisms. Table 1 compares a marine food chain with a terrestrial one.

All organisms in a community are functionally arranged according to their food source. Those obtaining food at the same relative position along the energy pathway are said to belong to the same trophic or energy level. Thus green plants, the producers, represent the first trophic level; plant eaters, the herbivores, the second;

Table 1
Marine and terrestrial food chain[1]

Tropical level	Marine food chain	Terrestrial food chain
Producers	Phytoplankton	Shrubs
Primary consumers (herbivores)	Herbivorous zooplankton	Deer
Secondary consumers (1) (general carnivores)	Anchovy	Tiger
Secondary consumers (2) (intermediate carnivores)	Herring	
Tertiary consumers (3) (top carnivores)	Tuna	

1. As is frequently the case, the marine food chain contains more steps than the terrestrial example. One explanation advanced for this common difference is that marine systems are older and have had a longer period in which to evolve.

Source: Arthur S. Boughey (ed.). *Fundamental Ecology*. San Francisco, Intext Educational Publishers, 1971, p. 23.

and so forth. (See Table 1.) A species population may occupy more than one trophic level, e.g. man, who is both an herbivore and a carnivore. In terms of trophic levels, the food chain illustrated above can be represented as follows (see also Table 2):

Autotrophs (producers) ⟶ herbivores (primary consumers) ⟶ carnivores (secondary and tertiary consumers) ⟶ decomposers.

Table 2
Trophic levels in ecosystems

Trophic levels	Gir Forest (India)	Savannah
Carnivores		Man
Carnivores	Man	Lions, leopards, cheetahs, hyenas, vultures, storks.
Carnivores	Lions, leopards, jungle cats, desert cats, hyena, jackals, foxes, mongooses, civets, snakes, frogs, lizards	Herons, flamingos, baboons
Herbivores	Buffaloes, sambars, nilgais, wild boars, four-horned antelopes, chinkaras, common langoors, hares, porcupines, mice, ants, termites, bees, butterflies, moths, etc.	Ostriches, elephants, hippos, giraffes, zebras, impalas, gnus, topis, fish, grasshoppers, other insects, crustaceans
Producers	Grasses, shrubs, trees	Palms, acacias, euphorbias, phytoplankton, grasses, shrubs

Sources: of 'Savannah' data: Helena Curtis. *Invitation to Biology*. Worth Publishers, Inc. 1972, p. 445;
of 'Gir Forest' data: Compiled from the data supplied by the ecologists of Gir Ecological Research Station.

In any ecosystem, and in the biosphere as a whole, the innumerable food chains are not isolated sequences but are interconnected in a complex pattern typical of a 'food web'. Generalized food webs operating in an estuary and a temperate deciduous forest are shown in Figs. 9(a) and (b). The species that comprise each trophic level differ from one ecosystem to another but the general pattern remains the same. That is, every food chain or web begins with autotrophs and terminates with decomposers—organisms of

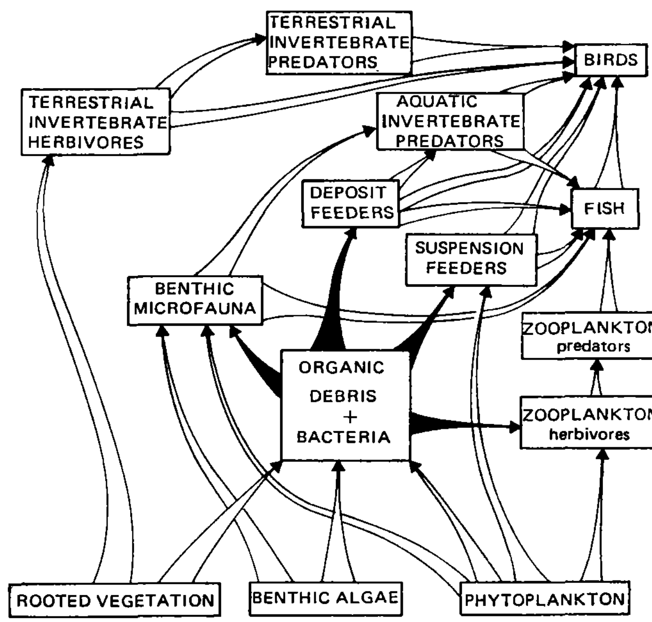


Fig. 9(a)
Estuarine food web showing food relationships of the main ecological groups of the estuarine animals

Source: J. Green. *The Biology of Estuarine Animals*. London, Sidgwick and Jackson, 1968.

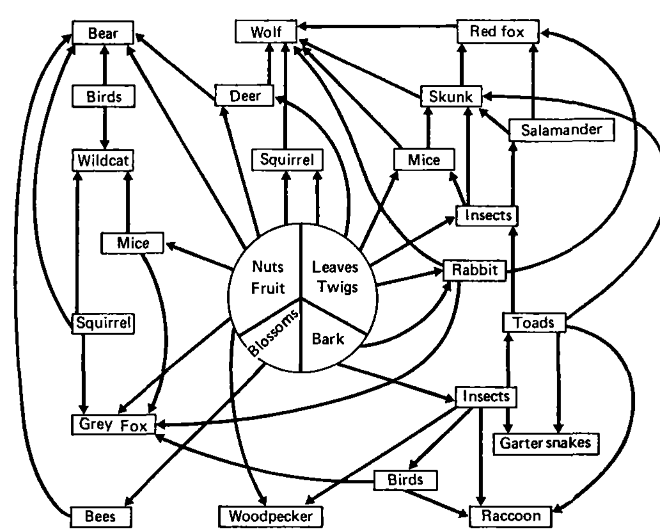


Fig. 9(b)
Food web in a temperate deciduous forest

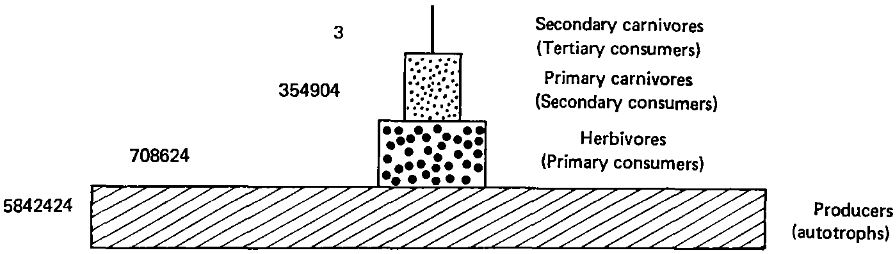
Source: J. W. Kimball. *Biology*. Reading, Addison Wesley Publishing Company, 1965.

decay, usually bacteria and fungi. The links between the producers and decomposers are variable. For example, the producers may die and be acted upon directly by decomposers in which case intermediate links are absent. The shorter the food chain or closer the terminal organism is to the beginning, the greater is the available energy that can be converted into living material and stored as food (the biomass) in addition to that which may be dissipated in respiration.

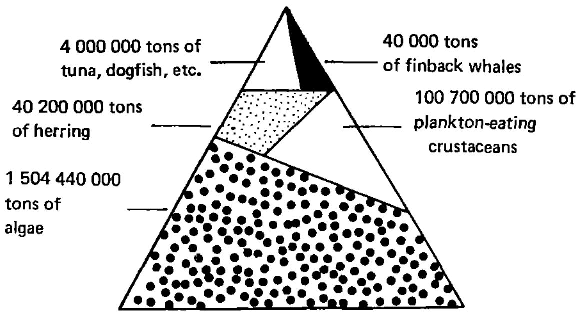
Ecological pyramids

Ecosystems in the biosphere have both structures and functions in terms of growth, relative abundance and the flow of energy. Such trophic structures and functions can be measured and represented graphically by means of an ecological pyramid—in which the first or producer level forms the base, the successive levels forming the apex being contributed by herbivores and carnivores in decreasing ratio. The following three general types of pyramid are known:

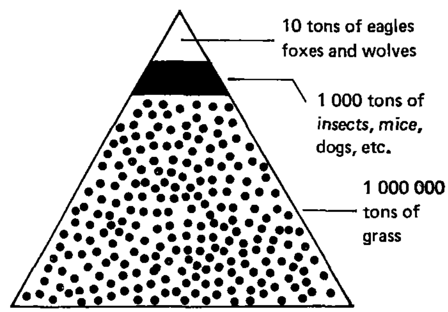
1. A pyramid of numbers depicts the total number of organisms ignoring the species composition. The major types of such pyramids in nature are shown in Fig. 10. There are difficulties in this approach since it equates a diatom with a tree (both producers) or a rabbit with an elephant (both herbivores).
2. A pyramid of biomass in which total dry weight (biomass) replaces the relative numbers of individuals. Nevertheless, it still has its limitations. Both pyramids indicate only quantity of organic material present at any time. The number and weight of organisms that can be supported at any trophic level depend on the rate at which food is produced at the level below, rather than on the amount of fixed potential energy at any particular time.
3. A pyramid of energy provides information both about the total amount of food and its rate of production at each trophic level. The 'number' and 'biomass' pyramids may be inverted or partly so, but a true pyramid is always upright. Moreover, the pyramid of energy is a necessary consequence of physical laws and is thus characteristic of all ecosystems.



PYRAMID OF NUMBERS



IN THE OCEANS



ON THE PLAINS

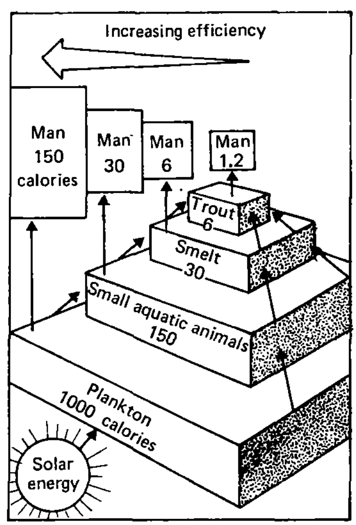
PYRAMIDS OF BIOMASS

Fig. 10
Ecological pyramids

Pyramid of numbers
(marine ecosystem)

Pyramids of biomass (marine
and freshwater spring ecosystems)
Source: L. J. Milne; M. Milne. *The Biotic World and Man*. Englewood Cliffs, New Jersey, Prentice Hall, 1965, p. 480.

Pyramid of energy (cold spring)
Source: L. C. Cole. 'The Ecosphere', *Scientific American* (New York), 1958, p. 198.



PYRAMID OF ENERGY

Productivity

The pyramid of energy brings us to the general concept of productivity. This is the function of the first trophic level and is defined as the rate at which energy is stored by photosynthetic and chemosynthetic activity of producer organisms (mainly green plants) in the form of organic materials which may be used as food by consumers at subsequent trophic levels—in any ecosystem or within the whole biosphere. The re-use of this food by subsequent levels is called assimilation. The productivity depends on the energy flow and is not the same as the standing crop or biomass at any given time represented by the first two pyramids. Productivity helps to maintain a certain biomass under certain conditions. For the earth as a whole, this mass is estimated to be about 410,000 million tons of protoplasm, of which about 290,000 million tons are of plant and microbial material and about 120,000 million tons are represented by animal protoplasm of all kinds.

The type of productivity may be differentiated, namely gross primary productivity, which indicates the total amount of organic substances resulting from photosynthesis in a given time and net productivity which is the amount of organic matter stored in plant tissues per unit time minus that utilized by them in respiration. In other words it represents new material added to the organism. The net production can be harvested and used to sustain man and the animals on which he feeds. Table 3 gives values of primary productivity for different ecosystems, in kc/m^2/day. It is noted that the primary productivity varies from one ecosystem to another. Since it is expressed on a real basis, low productivity does not necessarily mean that the plants themselves are less productive. It is just that there are fewer plants or some other factors are limiting. The efficiency of primary production is defined as the ratio between the flow of light energy stored by photosynthesis in carbohydrates and the flow of the light energy which the autotrophs intercept.

Table 3
Primary productivity in different ecosystems

Ecosystem	Primary productivity kc/m^2/day	Efficiency of primary production (percentage)
Natural ecosystems		
Subtropical	2.9	0.09
Desert	0.4	0.05
Arctic tundra	1.8	0.08
Coral reef	39–151	2.4
Tropical marine meadow	20–144	2.0
Tropical rainforest	131	3.5
Fertilized ecosystems		
Algal culture	72	3.0
Sugar-cane field	74	1.8
Tropical forest plantation	28	0.7
Water hyacinths	20–40	1.5

Source: H. T. Odum. *Environment, Power and Society.* New York, Wiley-Interscience, 1971, p. 83.

Pathways of matter

The flow of matter differs from the flow of energy in that matter is not lost to the system at each transformation. Thus matter in the biosphere tends to be used and re-used or cycled. This cycle is also referred to as a biogeochemical cycle or nutrient cycle. Two types of such cycles—gaseous and sedimentary—are recognized. Unlike the gaseous cycles, sedimentary cycles are essentially one-way flows. Elements like potassium, magnesium, calcium and phosphorus and other elements are released from the rocks on the land by weathering and then, following the movement of running water either in solution or in sediment, are carried to the sea where geological processes return them to storage in lithosphere. Since these elements do not form gaseous compounds at normal temperatures and pressures, they cannot return to the atmosphere and then to the land except by the birds or fish caught for food or by wind-borne sea salts scattered over land masses. The potential, then, exists for these elements to be in short supply.

In the gaseous cycle, however, the element or the compound may become converted to a gaseous form, diffuse through the atmosphere and thus get back over land or sea, to be re-used by the biosphere. The primary constituents of living matter—carbon, hydrogen, oxygen and nitrogen—all move through gaseous cycles which are discussed in detail elsewhere in the text. However, the major features of a material cycle are illustrated in Fig. 11.

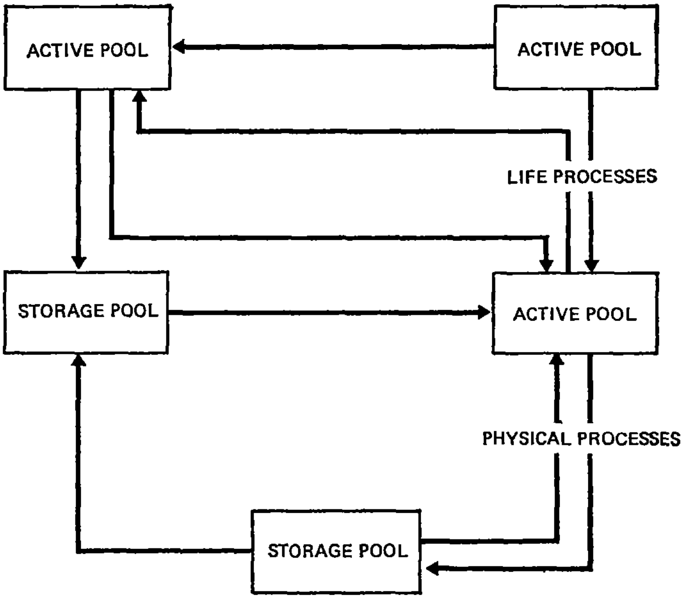


Fig. 11
General features of material cycle

Man's place in the biosphere

On the basis of ecological concepts and principles, man, like any other animal, exists as an intrinsic part of some ecosystem in the biosphere; though he may consider himself as apart from it. As such, his interactions with natural systems occur in the realms of physical and biological phenomena:

$$\begin{array}{ccc} & \text{Man} & \\ \uparrow\downarrow & & \downarrow\uparrow \\ \text{Abiotic component} & \rightleftarrows & \text{Biotic component}\end{array}$$

Like others in the biotic community, early man hunted and gathered food as well as reacting with the rest of the ecosystem for survival. But soon he developed tools, language and culture; and all these increased his impact on the surrounding environment. Discovery of fire and the wheel afforded him supremacy over the other members of the biotic community in the habitat. With further developments in his society, culture and civilization and by agriculture and industry, he assumed an ever greater role of a manipulator and an exploiter of the nature. With all this, he is still subjected to and affected by the environmental forces such as floods, hurricanes, earthquakes, landslides, etc. Many of the activities that we think of a peculiarly human have parallels in other animals, but the uniqueness seems to lie in his working on a different scale and in being a highly 'adaptable' component of the biosphere. The latter quality is not so much a factor of genetics as of his new and fancy ways to change the environment to satisfy his varied needs and desires.

When man modifies the environment to suit his own purposes (shelter, food raising, road building, etc.), he usually affects, and often destroys, habitats of other organisms. As a result, many species are currently threatened with possible extinction. For example, the pink-headed duck (*Rhodonessa caryophyllacea*) of India was confined to a swampy jungle habitat. It is now thought to be extinct primarily because of drainage of these swamp habitats, though its extinction was caused at least partly by excessive hunting. Adequate environmental planning and development must include an awareness of such effects of habitat destruction. Because of man's carelessness, ignorance or ruthlessness toward community structure, inadvertent influences on the environment often occur. For example, farmers sometimes make a special effort to eliminate a population of predators (e.g. the leopard or lion) in order to protect their domesticated animals (sheep or cattle), only to encounter a more severe pest (an overpopulation of antelope because of the removal of the predator).

Man is an omnivore, utilizing various trophic levels as his energy source; the particular choice being influenced by the wide range of cultural and individual differences due to ecological, social, religious and personal preferences. Likewise his other needs are influenced in the further exploitation of natural resources. Further, depending upon man's rationality and capacity to direct his action, the fate of ecosystems seems to fall under his activities and movements. In more ways than one, man improves and makes a further use of energy relations to serve his purposes but some-

times with the consequence of disturbing natural systems and bringing about a total imbalance in nature. In the end, man becomes helpless, powerless, and even more dependent on the forces of nature.

Man has developed sustaining and subsidizing mechanisms for increased energy yields. Highly productive areas of the world include man-made agricultural and crop lands. Increased production in these areas has been possible upon man's evaluation of the physical environment, his understanding of the principles of energy utilization and flow in ecosystems, the selection of crops that may best fit such conditions, and checks on the incidence of pests and other competing organisms.

However, in the process of man's utilization and exploitation of the ecosystems, several synergistic events have come up. For example, extraction, processing, and consumption of non-renewable natural resources bring about a pollution and degradation of our environment. This is nature's negative feedback that will have to be reckoned with sooner or later. Consequently, special attention must be given to such aspects including his carelessness, ignorance and ruthlessness toward ecosystem dynamics. Adequate environmental planning and development must, therefore, include an understanding of the structure and function of the specific ecosystem in the area being modified. What happens to the ecosphere will ultimately determine the fate of man himself. His needs of air, water and food, etc., all tie him intricately to his immediate environment.

Part I

Evolution of human populations

1 The evolution of man

Although man has frequently considered himself to be above, or somehow apart from, the natural world, he is finally beginning to understand that he is a small intrinsic part of it.

His needs are those of any other animal: water, food, a suitable climate, and shelter from predators. Beyond this, the problems he faces today are essentially of his own making, since he has found the means of satisfying these requirements in most cases. So well has he satisfied these needs, that he threatens their sources with his own success.

It is with this animal that we are concerned. Just what sort of an animal is he? Part of the answer to this question can be discerned by comparing man to his relatives.

Man among the animals

Quite obviously, man is a mammal, having hair and being nursed at the breast as a baby. He also bears great physical resemblance to certain other animals, such as the monkeys and apes. Undoubtedly, like any other animal he has some unique specializations, such as erect posture, opposable thumbs and unusual intelligence.

From the viewpoint of taxonomic studies we find that man is most closely related to the primates. Their anatomy and physiology show that they evolved from common ancestors.

The more remote of the living primates include the tree shrews, which closely resemble the insectivores, from whose ancestors the primates apparently arose. They are especially interesting in that they combine characteristics of both the insectivores and primates, but they exhibit modifications of the feet and brain that set them apart from the insectivores. The tree shrews are small nocturnal omnivores, found in South-East Asia.

Another group of living primates are the tarsiers, consisting of a single genus, *Tarsius*. They are also found in South-East Asia, but only in the islands. Their long

arms terminate in hands having slender fingers with large pads at the tips. These pads permit them to cling to smooth surfaces—an adaptation to their arboreal life. The enormous eyes of the tarsier mark him as nocturnal and are the most distinctive features of the little animal. They move with a peculiar hopping motion as they go about searching for insects, their principal food.

The lorises and their relatives, the bush babies, are found over much of Africa, India, and South-East Asia. Both have teeth that are typical of the prosimians and both are probably omnivorous, but the lorises seem to prefer insects more than do the bush babies or galagoes. The lorises are slow and deliberate in their movements, as opposed to the hopping, rapid movements of the bush babies. Both are nocturnal and arboreal.

The lemurs are found only in Madagascar, although there are several forms. There are few generalizations to be made about them, since they are variously diurnal, nocturnal, crepuscular, or indefinite in activity periods. Their size ranges from the mouse-sized (50 to 90 grammes) *Microcebus* to *Indri*, about the size of larger monkeys. They pose an interesting problem in evolution due to their unusual variability in a limited habitat. They have been isolated from the continent for a relatively long period of time without having recolonized it. The fossil record does not pre-date the late Pleistocene (Table 5).

In the suborder Anthropoidea are the remaining primates. Anatomical, physiological and biochemical evidence indicates that the monkeys are not as closely related to man as are the hominoid apes (Table 4). (See Plates 1 to 6, between pages 48 and 49.)

Table 4
A classification of the order primates[1]

Taxonomic designation	Common name
ORDER PRIMATES	
Suborder Prosimii	
Infraorder Tupaiiformes	Tree shrews[2]
Infraorder Tarsiiformes	Tarsiers
Infraorder Lorisiformes	Lorises and bush babies
Infraorder Lemuriformes	Lemurs
Suborder Anthropoidea	
Superfamily Ceboidea	New World monkeys
Superfamily Cercopithecoidea	Old World monkeys
Superfamily Hominoidea	
Family Hylobatidae	Gibbon and siamang
Family Pongidae	Chimpanzee, gorilla, orangutan
Family Hominidae	Man

1. Adapted from: J. Buettner-Janusch. *The Origins of Man: Physical Anthropology*. New York, John Wiley & Sons, Inc., 1966, 674 p.
2. Some authors think that these should be classified among the insectivores rather than among the primates.

The gibbons and siamangs occur in South-East Asia, along with the orangutan, while the chimpanzees and gorillas are indigenous to western and central Africa. Although the gibbon and the other primates are capable of bipedalism, the gorilla spends more time on the ground in this fashion than do the others. Yet even the gorilla

usually walks with the aid of his hands and cannot assume as completely an erect posture as man, due to the shape of his pelvis. All the primates are capable of grasping with the hands and most also can grasp with their feet, an ability essentially lost in man. With a truly opposable thumb he has, however, achieved the most control over a grasping ability. By this specialization his feet became more functional in walking with an erect posture and his hands were freed for the functions of carrying and manipulating. This, together with his increased intelligence, permits tool usage which allows creativity. In turn, these adaptations facilitated the development of speech, since his mouth was relieved from the function of carrying.

Pre-man

The evolution of modern man from his early ancestors is based on a very fragmentary fossil record and is subject to continual revision and controversy.

Table 5
An approximate geological time-scale for the Cenozoic era[1]

Epoch	Duration (millions of years)	Began (millions of years ago)
Recent	0.025	0.025
Pleistocene	1.75	1.8
Pliocene	11	13
Miocene	16	29
Oligocene	11	40
Eocene	19	59
Paleocene	17	76

1. Modified from: E. O. Dodson. *Evolution, Process and Product*. New York, Reinhold Publ. Corp., 1960, 352 p.

The earliest primates probably arose near the beginning of the tertiary period, in the middle of the Paleocene epoch, about 50 million years ago (Table 5). The Eocene primates that followed already showed a beginning of the rotation of the orbits toward the front of the skull, a characteristic that is typical of modern primates, including man. Another cranial feature that apparently started in the Eocene primates was the trend toward the present position of the *foramen magnum*, the large opening in the base of the skull through which the spinal nerve chord connects to the brain. This opening, located near the back of the skull in animals that walk on four feet, tends to be located under the skull in a more central position in those animals that hold their trunk and head erect. It was in the Eocene epoch that the Prosimii experienced their greatest adaptive radiation, producing several families including more than forty genera.

In the Oligocene epoch, beginning about 39 million years ago, when, strangely, the fossil evidence is less complete, the first of the anthropoids arose. The climate of the earth at this time was warmer than at present, resulting in the maximum spread of forests. The questionable ancestral form of the anthropoids, *Parapithecus*, is found first in this setting, probably evolving from prosimians. *Parapithecus*, along with

several other genera thought to be ancestral to Miocene apes, was discovered in the Fayum Depression in Egypt. As of this time, most of the Oligocene anthropoids are still known only from this area of the world, although a few fragments from Asia and North America are thought to belong to this group.

Also among the Oligocene forms is *Propliopithecus*, once thought to be ancestral to the gibbons, but now considered to be more closely related to, if not a direct ancestor of, the modern pongids and hominids. About the same time (Eocene?), the families Oreopithecidae, Pongidae and Hominidae diverged. One of these lines lead to *Oreopithecus*, a form present at the borderline between the Miocene and Pliocene epochs. Although it was likely not in the lineage of modern species, it did have an apparently erect posture and bipedal locomotion, which indicated that this hominid characteristic was developing at least 12 million years ago. Some investigators place *Oreopithecus* in the family Hominidae, but the relationship is doubtful. The family Pongidae includes many of the more difficult problems in hominoid taxonomy, perhaps because it includes our closest living relatives. This family probably separated early from the other lines soon after the establishment of the Hominoidea. The earliest representative known was *Aegyptopithecus*, found in Oligocene deposits in the Fayum Depression in Egypt. *Aegyptopithecus* or its close relatives led to *Dryopithecus* of the Miocene and Pliocene epochs.

The latter genus included the incorrectly assigned 'Proconsul', supposedly a progenitor of modern pongids. The dryopithecines were ape-like in appearance. In addition to giving rise to gorillas, chimpanzees and orangutans, they led to *Ramapithecus*, a hominid.

Ramapithecus is found in late Miocene and early Pliocene deposits. It was about

Table 6
A classification of Hominoidea, including probable human ancestors

Superfamily Hominoidea	Epoch
Family Oreopithecidae	
Apidium	Oligocene
Parapithecus	Oligocene
Oreopithecus	late Miocene
Family Hylobatidae	
Aeolopithecus	Oligocene
Pliopithecus	Miocene-Pliocene
Hylobates (gibbon)	Recent
Family Pongidae	
Subfamily Dryopithecinae	
Aegyptopithecus	Oligocene
Gigantopithecus	Pliocene
Dryopithecus	Miocene-Pliocene
Subfamily Ponginae	
Pan (chimpanzee and gorilla)	Recent
Pongo (orangutan)	Recent
Family Hominidae	
Propliopithecus	Oligocene
Ramapithecus	Miocene-Pliocene
Australopithecus	Pleistocene
Homo (man)	Pleistocene-Recent

the size of a gibbon, appreciably smaller than most of the dryopithecines, and probably walked upright, using its forelimbs to help in eating. Several skull and dentition characteristics lead us to believe that *Ramapithecus* was the progenitor of the Pleistocene hominids, most directly *Australopithecus*. (See Table 6 and Fig. 12.)

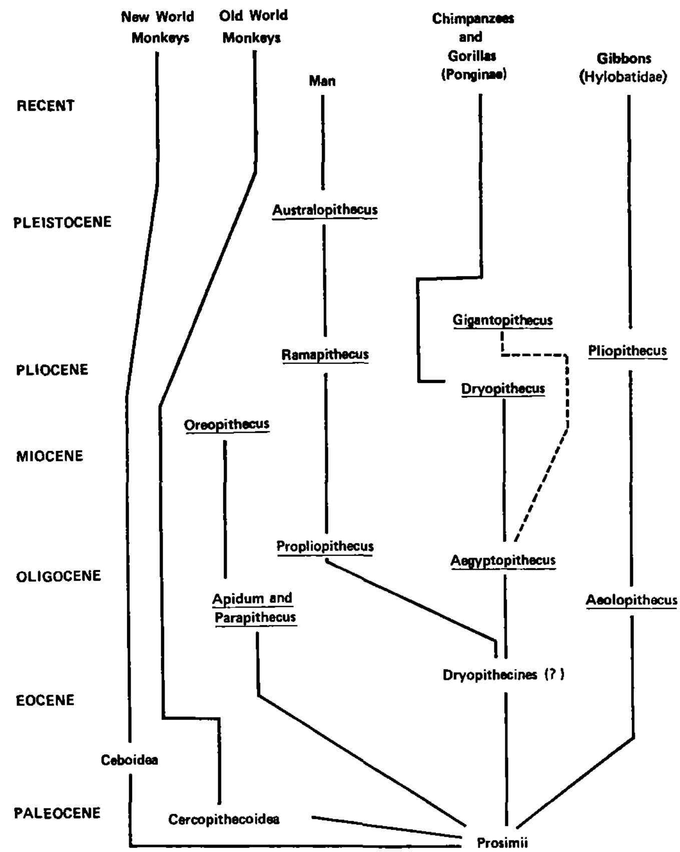


Fig. 12
A suggested sequence of evolution of the Superfamily Hominoidea

The origin of man

Up to the beginning of Pleistocene, the suggested lineage of man, then, follows this sequence. Prosimians of the Eocene epoch, a period of warm climate and much erosion, gave rise to the dryopithecines, ape-like animals of 50 million years ago. Through the spreading forests and warm climates of the Oligocene epoch and the cooler, volcanic Miocene, evolution continued, producing *Ramapithecus*, a gibbon-sized bipedal ape. In the drying, spreading savannahs, perhaps in Africa, 3 or 4 million years ago, emerged *Australopithecus africanus*, a tool-using pre-man.

There is some uncertainty about the dating of materials from Europe, Africa and Asia at the end of the Pliocene. The transition between the Pliocene and Pleistocene epochs was thought, until recently, to have been about 600,000 years ago. Within the

last decade, potassium-argon dating has established that, in East Africa, at least, this transition occurred about 1,750,000 years ago, which permits an appreciably longer period of time for the rise of man. This transition did not necessarily occur in all parts of the world at the same time. We know, for example, that what is called Pleistocene in Africa is a pluvial time, taken with some evidence as corresponding to the glacial times of Europe. The means of dating have improved a great deal recently, and undoubtedly further refinement of the time-scale will continue.

One observation we can make with some certainty is that, in the Olduvai Gorge in eastern Africa, where the date of the beginning of the Pleistocene was well established, *Australopithecus* was already present. It appears now that *Australopithecus* had appeared by 3 or 4 million years ago, evolving in the Pliocene savannahs and forests of Africa. The late Pliocene environment was progressively drier and warmer. Although, as a human habitat, it appears to have been rigorous, it was quite stable. With the spreading of the savannahs and the recession of forests, the forest apes must have become more and more restricted. An ape which found the means of subsisting on the savannah, however scant, had a large and growing habitat at his disposal.

Although there is still some controversy about it, it appears that there were two species of *Australopithecus* present at that time: *A. robustus*, the larger of the two at about 150 cm tall and 55 kg in weight; and *A. africanus*, about 120 cm tall and perhaps 40 kg in weight. It has been suggested that the larger were the males and the smaller were the females, but this hypothesis is not widely accepted. A more likely hypothesis is that two distinct species were present. This brings up the difficulties from an ecological point of view of explaining the presence of competing and concurrent species. It seems that one of them usually is eliminated. But the possibility still exists that the species were not really competitive. That is, if the larger species were a forest dweller and gatherer, and the smaller were a savannah predator, the competitive nature of the two species vanishes. Furthermore, small groups of these pre-men probably constituted mating bands as well. (See Plate 7.) Further isolation likely resulted from the sparseness of food on the savannah. It is unlikely that a predator, such as *Australopithecus africanus* apparently was, could have maintained a high density in such an environment.

Thus, genetic drift could have been as important as natural selection in the early differentiation.

There are many other interesting inferences to be made about pre-human evolution. One of the more plausible, which helps to explain the development of *A. africanus* concerns the comparative difficulty of life in the forest and in the savannah. In the forests, the home of the apes, there is an abundance of food and shelter, with little challenge to ingenuity. Natural selection would not place much favour on brain development except for those basic primate characteristics of increased visual acuity and stereoscopic vision, which would be helpful in an arboreal existence (to avoid frequent falling). Colour vision would be helpful in fruit and other food identification. If you take these same basic primate senses and apply them to hunting in a place where there are few animals, you find that they would be helpful. In addition, however, there would be a distinct advantage for those individuals or populations

which could make use of innovative hunting techniques. In other words, the problem posed by the harsh environment would play a major role in brain development. This is coupled, also, to the advantage of using tools. It seems impossible to say which development precedes which, but it appears that a relationship existed between the erect posture (permitting wider use of the hands), brain development, and tool use, that would favour the interaction and extension of all three. In any event, these reached a stage of development in *A. africanus* which far exceeded that of their ancestors.

Evidence is lacking to show a strong correlation between tool use and cranial capacity, but *A. africanus* did have prepared tools (Fig. 13). These are to be dis-

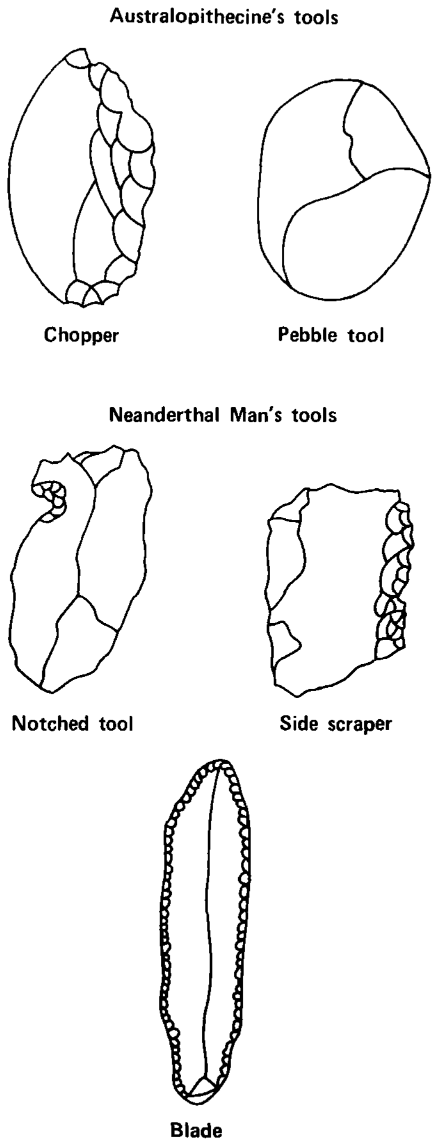


Fig. 13
Some representative tools of pre-man and early man

tinguished from casually grasped tools, such as a branch of a tree or a bone that a chimpanzee might pick up, use for a specific purpose, then discard. Rather, the prepared tool does not exist upon discovery in the form in which it is to be used. Then the user must perceive a changed form before the preparation and, if the tools are to be made and used by others, symbolic thought and communication are expressed by the users. This is not a consequence of tool-making, but a prerequisite.

In human lineage, stone tools first appear to be associated with the australopithecines, but it should be pointed out that these are the ones that would be preserved, even though they are not likely the first tools to be prepared or used. That is, wooden tools would have been easier to fashion, but would not be preserved. Tool-making and use must be older than our fossil evidence. The australopithecines made and used only 'pebble tools', as far as the fossil evidence indicates. These are merely stones that have a few flakes knocked off one end to provide a sharpened edge, which may be used in a chopping fashion while the round end is held in the hand. They are clearly fashioned tools and not accidents of nature because of the number that are found associated with fossil australopithecines. They were used, essentially unchanged, until the beginning of the middle Pleistocene. With present-day rates of change, it is difficult to comprehend a situation wherein one kind of tool, probably the only tool or one of few, could have been used unchanged for millions of years, but that is apparently the case.

It appears now that *A. robustus* became extinct without leaving descendents, but the hunting mode of life that was developed in *A. africanus* was successful. From the fossilized bones of his kills it was found that he subsisted on small animals, such as rodents and birds. In the Pleistocene, the later deposits show that his descendents hunted progressively bigger game, such as deer, bears, and bison.

At this juncture, it is necessary to recall that the foregoing prehistoric development is tentative. Indeed, a recent find by Richard Leakey at Lake Rudolph in eastern Africa casts at least some doubt on the reported chronology. The find consists of the skull of an hominid with a cranial capacity in excess of 800 cc from deposits older than 2.8 million years. A highly specialized small-brained *Australopithecus* has been found in the same environment at that time period. Further information is certain to develop in this field. A transitional site was discovered in a cave in France, the earliest known human living site outside Africa. It may be australopithecine and includes the use of fire, pebble tools, and larger prey animals than is found in African australopithecine sites.

What caused man (or pre-man) to migrate from Africa and spread over Europe and Asia is open to question, but the movement corresponds to the upheavals, in lower- and mid-Pleistocene times, which led to the Great Rift Valley. It has been suggested that this changed the course of major rivers from east-west to north-south directions, removing barriers. Also, the glaciation in northern climates resulted in cooler, wetter weather in Africa, probably changing the character of northern deserts and making them passable. But the fossil record indicates that by mid-Pleistocene times, perhaps more than half a million years ago, true man had evolved. These early men are known by several different names (pithecanthropines, Java man, Peking man,

and others), but are probably most conveniently called the species *Homo erectus*. But the variety of names does not conceal the agreement that he was a true man.

The world of early man

Homo erectus is described as having massive brow ridges, the *foramen magnum* further forward under the skull, straight leg bones, wider hips, rather large (but hominid) teeth and jaws, and no chin. His legs and arms were proportionately similar to modern man. Whereas the australopithecines had a cranial capacity (brain size) of around 600 cc, Java man's was about 775–900 cc and Peking man's was about 850–1,300. This compares with modern man's cranial capacity of about 1,350 cc.

Peking man used fire, made a variety of tools, and probably cooked food, since charred bones of deer were found in his living site. *Homo erectus* had changed the hunting habits of his predecessors by going to larger and larger game and was capable of killing animals much larger than himself. This was probably due not only to the use of weapons, but also a higher degree of organization in the hunt. His prey included elephants (which even the big cats did not kill, apparently), horses, deer, oxen, and rhinoceroses, far too large to be killed regularly by a lone hunter. This supports the premise that co-operation was needed.

Homo erectus had survived the comparatively unstable climates of the Pleistocene and, as the evidence indicates, was very widespread. His co-operative mode of living was successful. But he was succeeded in a later interglacial period by a larger species, *Homo sapiens*, who kept his predecessor's abilities and improved on them.

A transitional form, but one clearly belonging to the later species, was discovered in an English gravel pit amid evidence of favourable climate for man (game was apparently abundant for the emerging *Homo sapiens*). This early transitional form was named Swanscombe man. Although his cranial features place him in *Homo sapiens*, he did have a more sloping forehead and lower cranial vault than modern man. The interglacial deposits in which Swanscombe man (and later finds of Steinheim man) were found date back about 250,000 years, but the tools used by these early men were essentially unchanged from those of late australopithecines, except that they were more carefully manufactured. The return of glaciation soon after this period left the fossil record almost blank for the next 150,000 years, until about 100,000 years ago.

With the recession of the glaciers a new man appears in the fossil record. The paucity of evidence for the glacial time during which he developed poses a number of questions. For example, the new man appears quite startlingly changed for the short geological time for the changes. Furthermore, in some respects, they were more primitive than Swanscombe and Steinheim men. They had heavier brow ridges and lower cranial vaults and their bodies were short and stocky. These were the Neanderthal men, the best known of the cave dwellers (Plate 10).

Neanderthal man could be considered a race or subspecies of *Homo sapiens* due to the cranial features. He likely walked as erect, or nearly so, as we do and, by virtue of this and other features, was a man. His body type was well adapted to a cold

environment, minimizing heat loss through shortened extremities and thick bodies, much as is the case in modern Eskimoes. His passage out of his homeland was at least partially blocked by glaciers, so his adaptations physically and physiologically would permit him to survive without migrating out. Such a degree of isolation can account for his differentiation. It appears that his survival during the coldest periods, when hunting was very limited and his movements through the snow even more limited, would be extremely unlikely or impossible without food storage. But he was a big game hunter, killing very large animals, and cold storage was abundant. He used fire extensively, but a supply of fuel may have been difficult to maintain. He probably burned the fats of the animals he killed.

An even more human aspect of Neanderthal man is the presence of several records of his ceremonial burials of his dead and of other unknown rituals. His social amenities were well developed. The rituals frequently involved the remains of cave bears, perhaps related to his respect for them. After all, in order to occupy a cave, he probably had to drive out the bears occupying them. These rituals, particularly the burial ceremonies (incidentally even involving gathered flowers), imply that he had acquired a respect for the individual, unknown in any earlier men.

But Neanderthal man, in spite of his advanced culture and elaborate life style, did not survive. A number of suggestions have been made to account for his disappearance about 35,000 to 40,000 years ago. One account indicates that he may have been systematically destroyed by his successor, Cro-Magnon man, but this possibility seems extremely remote. A more likely explanation involves the peculiar adaptations of Neanderthal man, who flourished in the colder periods. His disappearance coincides with the retreat of the last Pleistocene glaciation in Europe and the invasion of the region by Cro-Magnon man. If he was well adapted to a cold climate and faced with a warmer one, he could have found it difficult to adapt. If, further, he was faced with a new competitor (a big game hunter like himself, but more efficient), which he was in the form of Cro-Magnon man, it is possible that he would succumb. Many other species have become extinct in what appear to be similar circumstances.

Long before Neanderthal man disappeared, other populations were developing in the eastern Mediterranean region. These were the Cro-Magnon men, much more like modern man in physical aspect, lacking the heavy brow ridges and with a cranial shape and face much like present man. Their life style was similar to that of Neanderthal man, except that they tended to form co-operative associations between bands, probably sometimes massing several hundred people together for certain events. The term tribe could probably be applied here.

Their co-operative ventures apparently included mass hunts, where large numbers of big game animals were killed at one time. This is in contrast to the Neanderthal band of hunters, of perhaps ten to twenty men, for whom killing a bear or elephant was a considerable accomplishment. The bands were still existent in Cro-Magnon man, but it was the co-operation between them that permitted large-scale ventures.

The cohesiveness of the tribe likely included matings between bands which would enlarge the genetic population to a point where genetic drift would be unimportant. It has also been suggested that warfare could have arisen in this social setting. There is

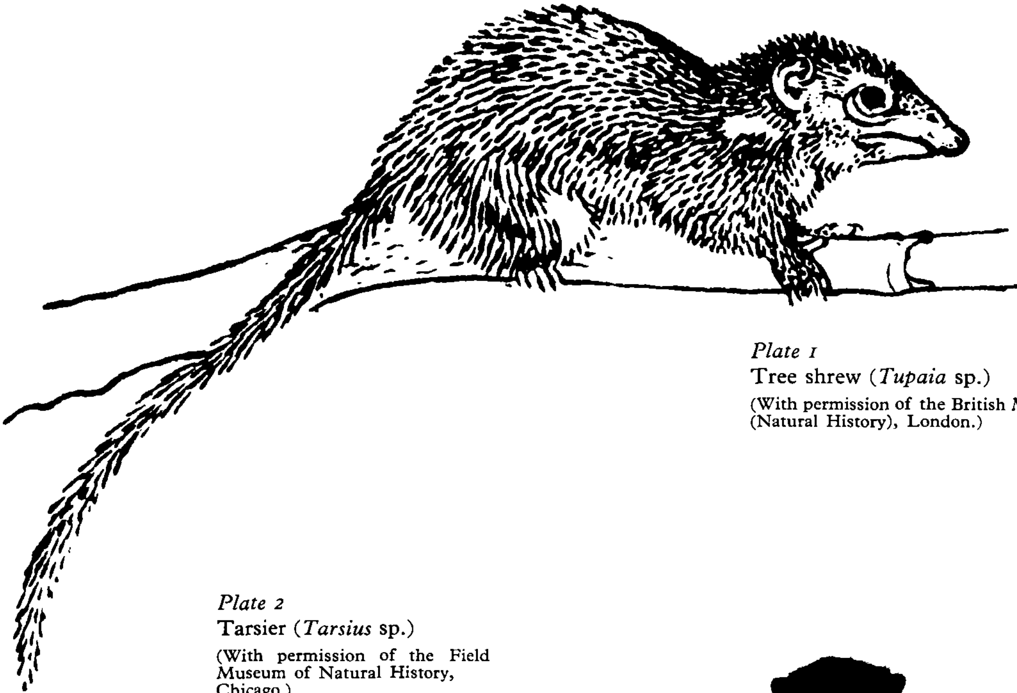


Plate 1
Tree shrew (*Tupaia* sp.)
(With permission of the British Museum (Natural History), London.)

Plate 2
Tarsier (*Tarsius* sp.)
(With permission of the Field Museum of Natural History, Chicago.)

Plate 3
Lemur
(With permission of the Field Museum of Natural History, Chicago.)

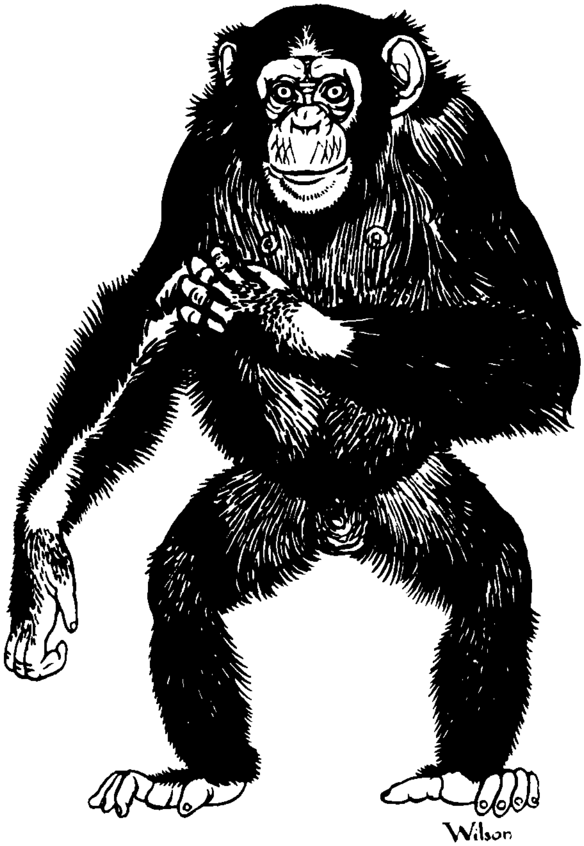


Plate 4
Chimpanzee (*Pan* sp.)
(With permission of the British Museum (Natural History), London.)

Plate 5
Orangutan (female)
(With permission of the Field Museum of Natural History, Chicago.)

Plate 6
Gorilla
(With permission of the Field Museum of Natural History, Chicago.)

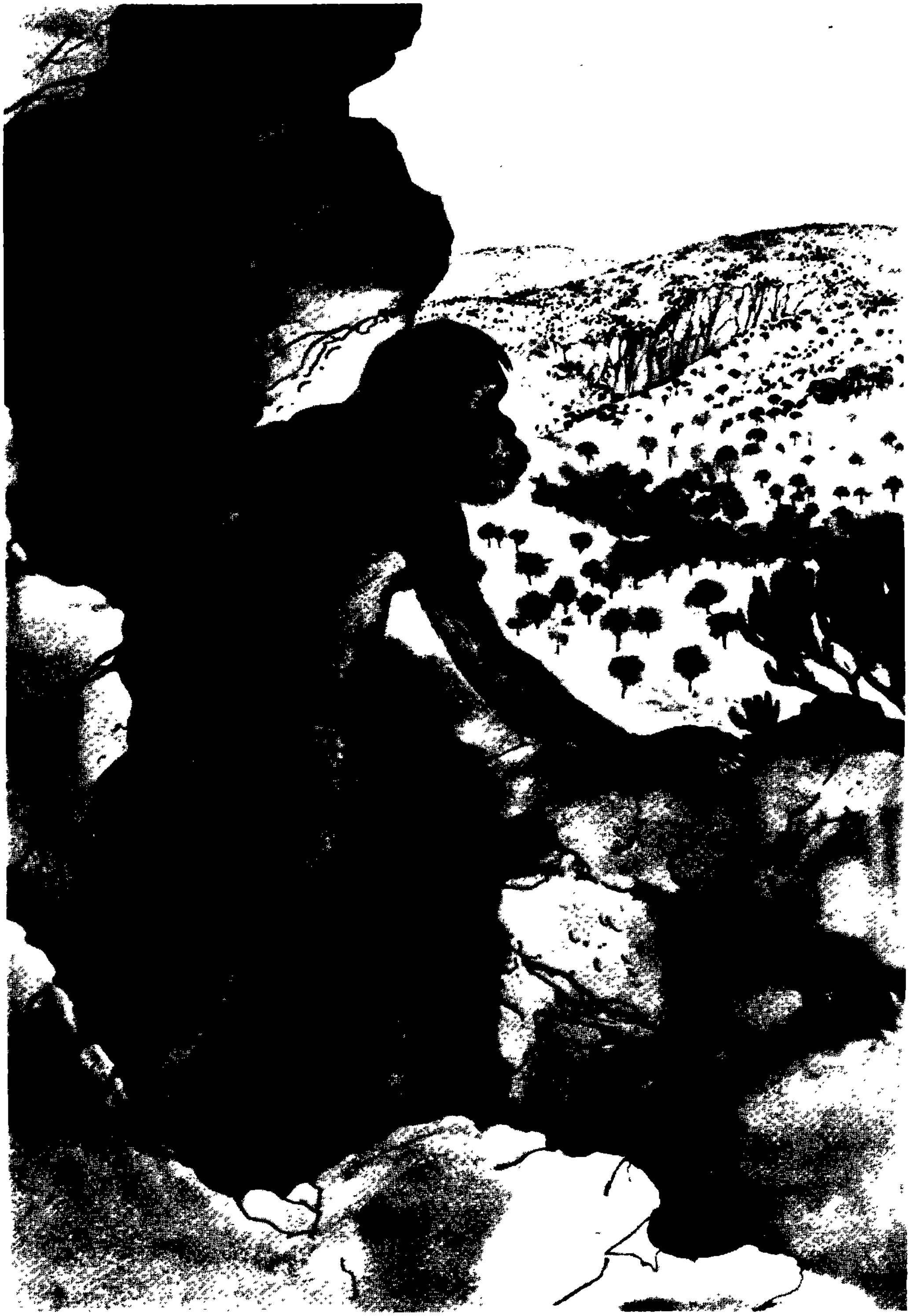
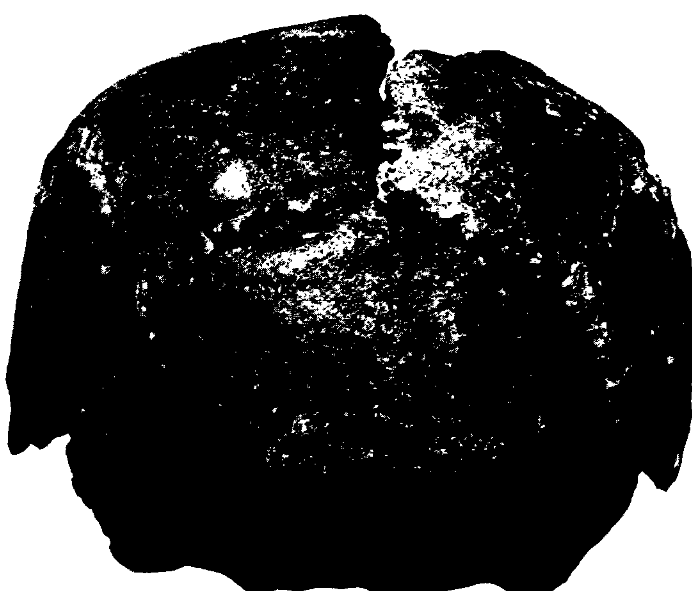


Plate 7
An artist's conception
of *Australopithecus africanus*,
at Makapan, South Africa

(With permission of the British Museum
(Natural History), London.)

Plate 8 – A
Side view of Swanscombe skull (*Homo sapiens*, early form), showing left parietal and occipital. Swanscombe is in Kent, England. The discovery was made in 1935 and 1936

Plate 8 – B
Swanscombe skull from back
(With permission of the British Museum (Natural History), London.)



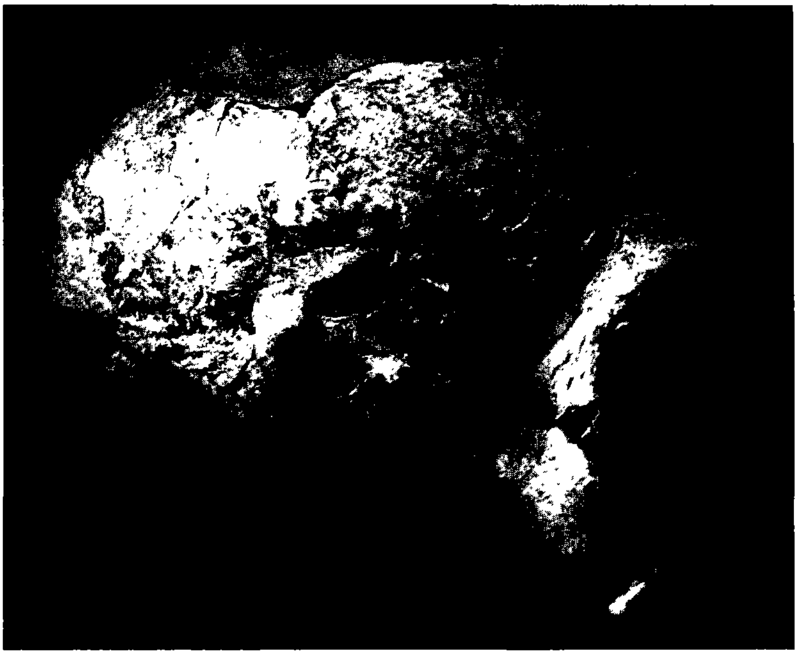


Plate 9 – A
Side view of Gibraltar skull
(*Homo sapiens neanderthalensis*)

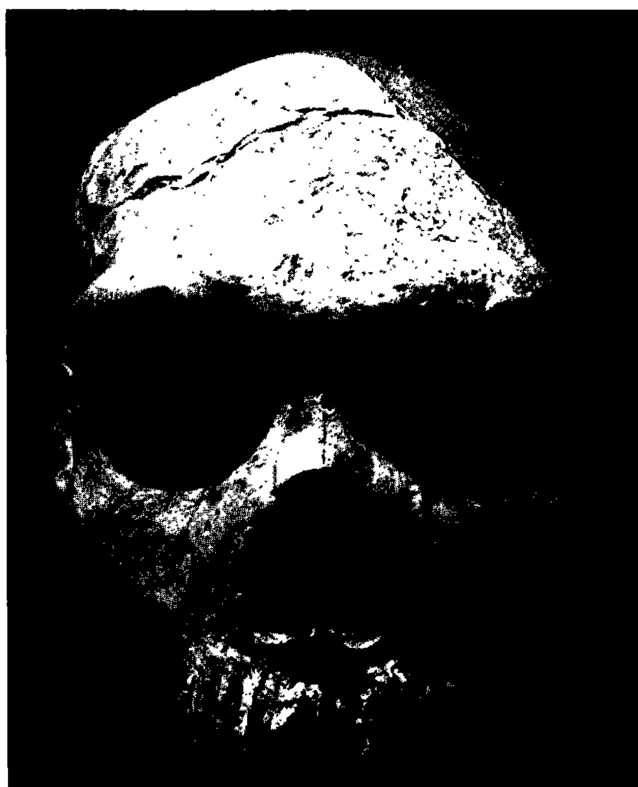


Plate 9 – B
Front view of Gibraltar skull
(With permission of the British Museum (Natural History), London.)

Plate 10
An artist's conception of *Homo sapiens neanderthalensis*
(With permission of the Field Museum of Natural History, Chicago.)

Plate 11
An achondroplastic dwarf showing the abnormal limb development

evidence that Neanderthals suddenly disappeared when confronted with Cro-Magnon man. But in other sites, it appears from the tool assemblages that Neanderthals may have been assimilated culturally and from skeletal remnants, that the two types may have interbred. In any event, when competition for food or living space occurred, Cro-Magnon man won.

Cro-Magnon man left paintings on the walls of caves where he lived. These show considerable ability, attesting to the high degree of abstraction he possessed. Though Neanderthal man showed crude forms of adornment (as of the graves) and probably also used pigments on his body, Cro-Magnon man expressed a truly artistic sense of proportion in his works. The brain of Cro-Magnon man had already reached the size of modern man's brain, so we may say that cranial capacity has not changed much since then. This does not imply that change has ceased, since measurable change would not likely appear in this brief time span. But the pattern of selection was clearer in his simpler existence that it is in ours. With an improved brain could come the improvement of existing skills, which might involve not only the cerebral cortex (intelligence or general ability to learn), but also the cerebellum (improving manipulative skills). This could have led to the invention of new tools, permitting preferential survival of the population to which the inventor belonged. The expanding population could then spread over adjacent areas, increasing the likelihood of hybridization, thereby enhancing variability (including the possibility of improved brains). In addition, colonization of these new areas could lead to new activities and new bases of selection, which might also lead to a higher proportion of improved brains in the population.

Implied in this selection cycle is migration, which was characteristic of *Homo erectus*, while before, in the large savannahs of the Pliocene, movement of australopithecines would not lead to much chance in their environment. Thus, it would not have been until the Pleistocene movements that the cycle would be very effective.

The development of different tool associations, probably reflecting different cultures of Cro-Magnon man, is evident in the fossil record. Although the following account is primarily European, this reflects only that we have not yet collected enough data from other areas. Beginning about 35,000 years ago and continuing for about 15,000 years, there were two known cultures: the Perigordian, which probably developed in western Europe, and the Aurignacian, which probably developed elsewhere (eastern Mediterranean (?)) and also occupied western Europe during the same period as the Perigordian. It was in the Aurignacian culture that the cave paintings were originated and brought to a high degree of sophistication. How these cultures coexisted for some 15,000 years without blending is not known, but there may have been some differences in the kinds of terrain they inhabited.

These cultures disappeared about the same time the Solutrean culture arose, 20,000 years ago. The Solutreans brought tool-making to a new state, i.e. in which the functional aspects of the tools were superseded by their artistic shapes. Some of the knives were beautifully shaped, but so slender that they could not have been used without breaking. Also, in some sites, the number of such tools of a particular kind suggest that they may have been made for export or trade. There is also indirect

evidence in the form of apparent arrowheads that these people had, or may have invented, the bow and arrow.

The Solutrean culture, a brief one, was succeeded by the Magdalenians about 17,000 years ago. Their tools were made from bone, antlers of reindeer and ivory, and were elaborately decorated with engravings. Most of the engravings were of characteristic animals of the time (including some of men), but there were also abstract designs. Their weapons included elaborately devised harpoons, which presumably were used in the mass hunts for reindeer, one of their staple foods. The records of the Magdalenians indicate that they became well established and flourished. Suggestions as to why this was true are varied, but include the change in the climate at about this time and possible cultural innovations.

The glaciers were receding, raising the ocean levels above the edge of the continental shelf. This greatly increased the productivity of the ocean, which is greatest in the shallows. Some of the resulting increase in marine life would have become available even to men living inland, e.g. increased (perhaps many-fold) spawning runs of salmon. The region inhabited by man was also on the route of the migrating birds which moved north to use the breeding grounds exposed by the retreating glaciers. The effects of both these factors would have been to provide food in periods when the reindeer were less available, helping to establish more permanent settlements and the expansion of existing ones. This did happen, as the fossil record shows that in the later phases of their culture, 12,000 to 14,000 years ago, their population increased markedly. The estimates are necessarily weak, but are of some comparative value for the population sizes of these men, but they are estimated to have increased in France from a population of 15,000 to 50,000 during this time.

As the glaciers retreated further to the north, so also did the climate and flora of their range change. The food of the reindeer 'followed' the retreating glaciers and the reindeer followed their food. A portion of the people followed the reindeer, while a portion of them remained behind, changing their life style to adapt to the new conditions. These were the Azilians.

The Azilians were confronted with less desirable conditions than had existed for the Magdalenians. Refining the hunting of smaller game became part of their lives, and they, as did other peoples of the world at that time, experimented with the cultivation of food plants. Their diet, more varied, was probably a better diet than the big-game hunters had. The success of this venture is supported by the evidence of rapid population increases in these areas, particularly the eastern Mediterranean region. The increase in population pressure may have led to the transplanting of food plants from their native habitat to occupation areas nearby. The next step, at about the same time, was to domesticate animals, such as sheep and goats. There is some evidence that this took place over 8,000 years ago, perhaps as early as 10,000 years ago.

These changes, and the ones that followed, show much more rapid sequences in Cro-Magnon man than occurred in his predecessors. This is because the changes became predominantly cultural rather than genetic. Furthermore, cultural changes are intentional and directional, a new feature for man. This may be compared to the behaviour of a new and beneficial genetic mutation. That is, when such a gene appears,

it offers a selection advantage to the individual that expresses it. Hence, the gene tends to replace its alternative alleles. The pronounced difference between cultural innovation and genetic mutation is that the cultural innovation may be disseminated in an extremely rapid manner, while genetic changes may replace existing alternatives only over many generations.

This, however, is related to a phenomenon of interest to modern man: over-population. The term 'over-population' has meaning, in an ecological sense, only in the context of the environment. A given, finite environment, due to non-biological conditions and the state of the biotic community living there, is capable of sustaining a limited population size for any particular species. This presumes that that species does things in a definite, constant way, such as eating the same foods, defending the same territorial size, etc. This appears to have happened for the australopithecines, at least through the latter part of the Pliocene epoch. Once he reached an equilibrium state, consonant with the conditions of his environment, his population density became stable or comparatively so. Provided neither he nor the environment changed markedly, he could continue to follow his life style indefinitely. But in the Pleistocene, his environment began to change, although not as drastically as it did in Europe. It is not surprising that evolution produced a better-adapted animal than he, probably by merely changing him, and that he did not survive, in his earlier form, these changes. But man evolving through *Homo erectus* to *Homo sapiens*, began adapting in a new and different way: genetic culturally. Especially in Cro-Magnon man, one can see the rate of change being overtaken by more rapid cultural change which, in the Azilians, led to intentional modification of the environment. This was done in such a way that rapid population increase resulted. The population increase is now heavily dependent on the continuation of these cultural patterns. So long as these patterns and their changes remain beneficial, the population increases might be sustained. But clearly, the species has created for itself a disequilibrium in terms of rates of changes. If an error should be made, or if cultural change should fail to keep pace with the increasing population, the environment could no longer support the species in the earlier fashion. Thus, unless the species numbers come into equilibrium with the new circumstances, excess numbers will be produced, which cannot be sustained; hence, over-population.

Looking back in time, then, we can see these elementary mechanisms emerging in the Azilians, but not becoming serious problems until modern times. Thus, man changed from primarily hunting and gathering to agriculture and, soon after, entered his period of recorded history.

2 Genetics background

Transmission of characters

It is now widely known that chromosomes in the cell contain genes which provide information for building new living units (cells, organisms). A particular bit of information (a gene) controls or influences one or more traits and is located in a specific site (locus) on a particular kind of chromosome. There are several kinds of chromosomes in each cell and, in the more complex animals such as man, except for the reproductive cells (gametes), there are two of each kind. In man, there are normally twenty-two pairs of chromosomes (autosomes) and two sex chromosomes (gonosomes), which are of the same kind in the female (the two X chromosomes) but are different in the male, who has an X and a much smaller Y chromosome.

It then follows that, for a particular characteristic of the individual, in the simplest case, there are two genes for the same locus in each cell, one on each member of the chromosome pair involved. Although the genes occupying that locus affect primarily only one characteristic, say one aspect of blood chemistry, it may be seen that differences exist in genes of that basic type, producing variation in that particular aspect of blood chemistry. These different genes that can occupy the same locus are alleles (antagonistic) of each other.

This is well illustrated by the *MN* blood group system. There are only two different alleles controlling the system: *M* and *N*. Since each cell has two loci for these genes, and since all the body cells of an individual are alike in this respect, we may find three possible types of individual: the homozygotes, who have both alleles alike (*MM* or *NN*), and the heterozygotes, who possess both types (*MN*).

The individual of the *MM* type has antigens *M* in his red blood cells, while the person homozygous for the *N* gene (having genotype *NN*) will then produce the *N* type antigen, and the heterozygous person (*MN*) produces both types of antigens.

But it is the transmission of these traits from one cell to an offspring cell or from people to their children with which we are concerned from the standpoint of genetic counselling or evolution. Basically, the process consists of replicating the genes and then dividing them in an equitable fashion among daughter cells. Since the genes are located on chromosomes, the transmission of traits is a matter of making essentially

exact copies of the chromosomes and then dividing the copies and the originals between the daughter cells by mitosis. Of course, in the production of reproductive cells which are destined to unite with other cells, the process is meiosis.

Prior to the visible process of cell division, the nuclear material undergoes a series of changes, some of which are as yet not well known. Chemically, the gene is constituted by DNA, a compound having the form of a double helix. The side chains are composed of alternating structures of deoxyribose sugars and phosphates, the two side chains being connected by pairs of four different kinds of molecules: adenine, guanine, cytosine and thymine. The sequences of these 'connectors' constitute the genetic information. Since adenine (*A*) can only bond with thymine (*T*), and guanine (*G*) only with cytosine (*C*), the mechanism for replication is as seen in Figs. 14 and 15. The chain breaks between each pair of 'connectors', thus exposing a site for recombination. Nutrient materials present in the nucleus then reconstitute the missing half of each chain, resulting in two complete and identical sets of instructions.

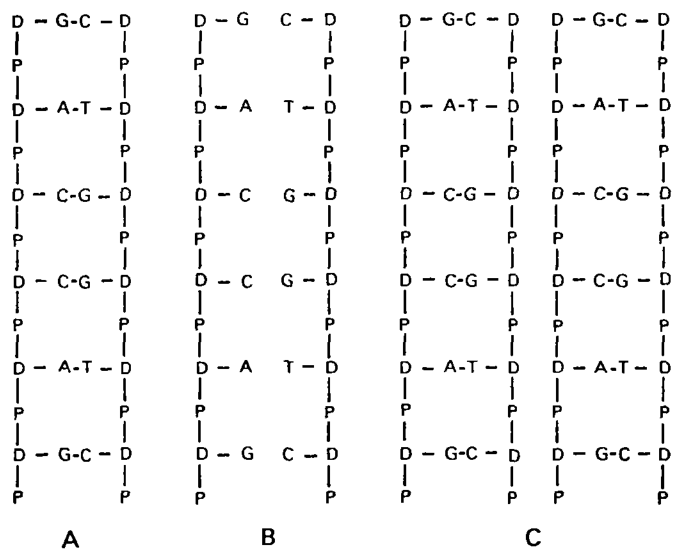


Fig. 14
A. Complete portion of DNA. B. Two DNA halves uncoil and separate. C. Each half forms a complement resulting in two molecules of DNA exactly like the original

In any event, it appears that one of each of these chains is a daughter gene and the two sets of genes form the two chromatids, which become apparent, at metaphase, as the halves, joined together in pairs, which form the chromosomes. The centromeres which join the chromatids become aligned along the future division plane of the cell and become connected to strands of cytoplasm (microtubules) which pass to opposite poles of the cell. Then the two members are separated at their only remaining connexion, the centromere, and are apparently drawn to opposite poles by the strands. Following this, a new nuclear membrane appears around the chromosome groups and the cell divides between them, resulting in two cells with identical chromosome complements. By this process, continually being repeated, the body of the individual is elaborated from the zygote (fertilized egg). Early in the development of the body,

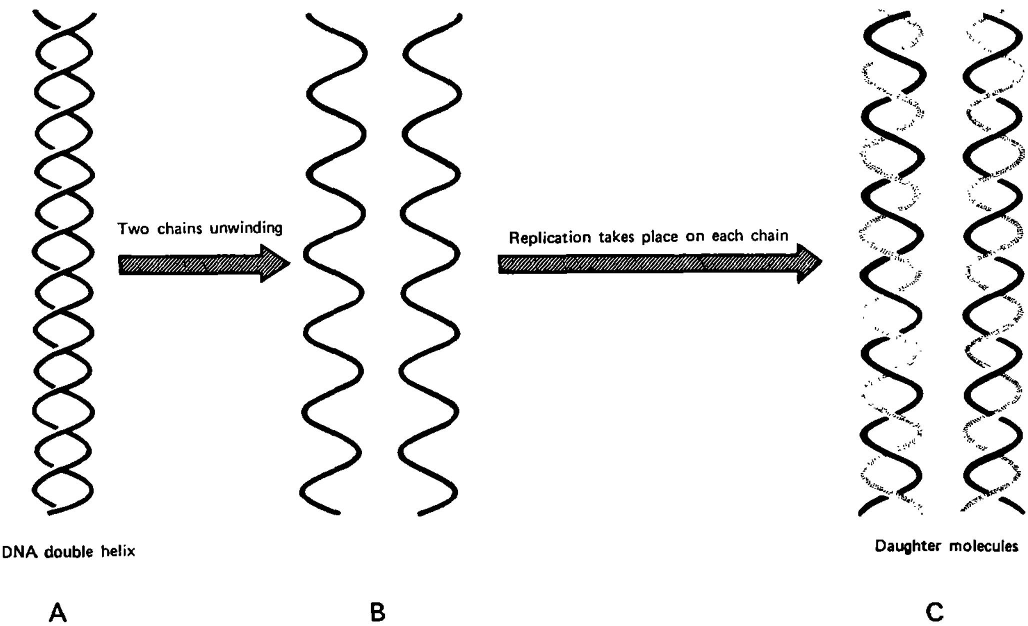


Fig. 15
A. Two chains of DNA are coiled about one another forming a double helix. B. The two chains unwind as replication begins. C. Each chain replicates a complement resulting in two daughter molecules

some of the cells are set aside, later to be used in reproduction. They will become incorporated into the gonads of the individual, where they will undergo meiosis. This reduction division complemented by fertilization allows for the evolutionary advantage of recombining the genes into a vastly more variable array of possible genotypes than is possible with asexual reproduction, and doing it in a shorter period of time than mutation could provide.

In meiosis, one replication of the chromosomal material is accompanied by two cell divisions, rather than just one. Thus, the pairs of chromosomes are separated, leaving each cell with only one member of each pair. These gametes are then able to combine with other such cells and produce a zygote with the normal body cell number of chromosomes, the complete set. The meiotic process, depicted in simplified fashion in Fig. 16, begins with the replication of the genetic material during the interphase (as a comparison, see Fig. 18 for mitosis). After the gradual condensing of the replicated genetic material, as in mitosis, the nuclear membrane disappears, the spindle forms, and the centromeres line up on the plane of future division. At this point, the two chromosomes of each type are paired, and each consists of two chromatids. The chromosomes then separate, each consisting of two chromatids (Fig. 17). Following cell division, the centromeres again line up as before and the two chromatids of each

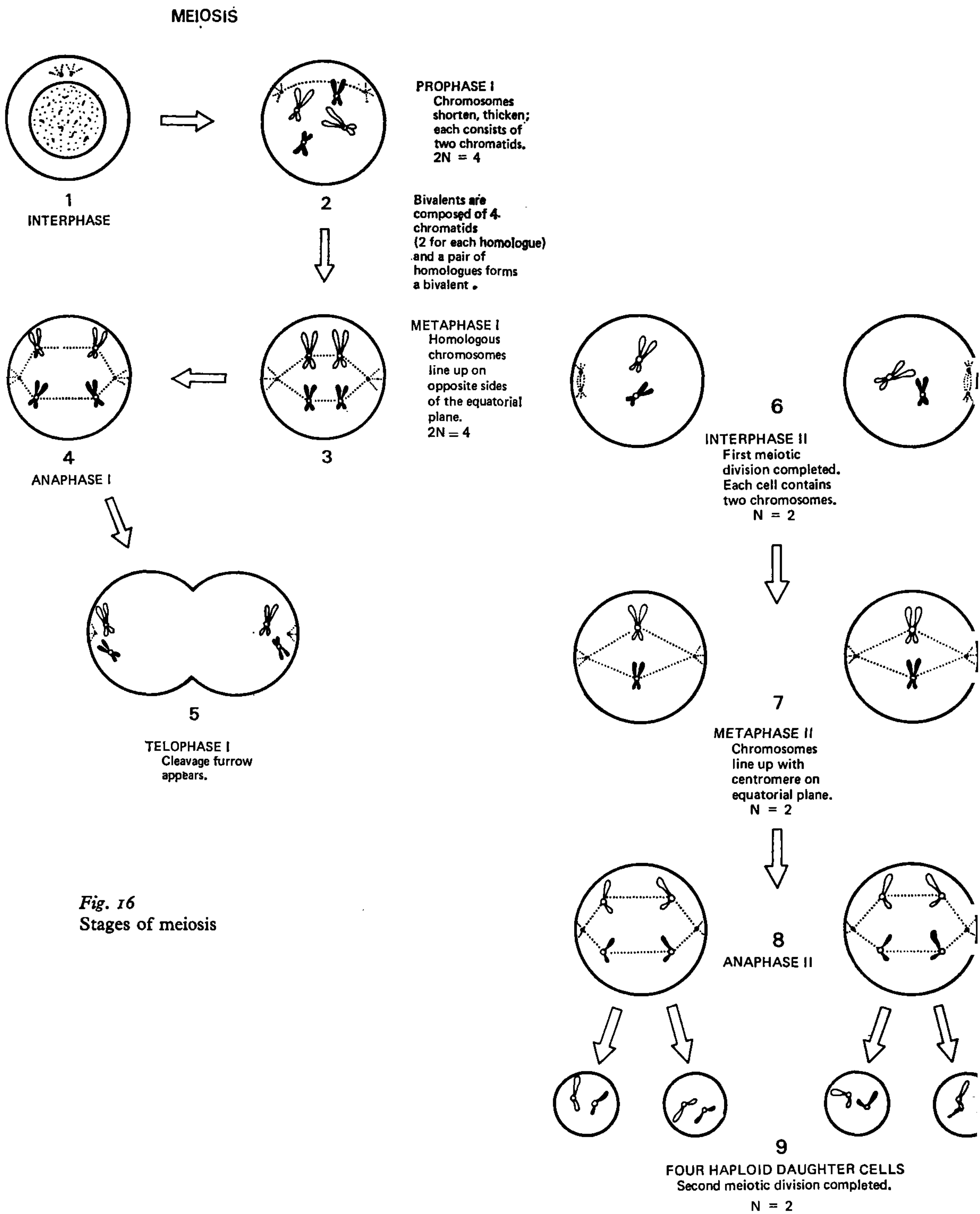


Fig. 16
Stages of meiosis

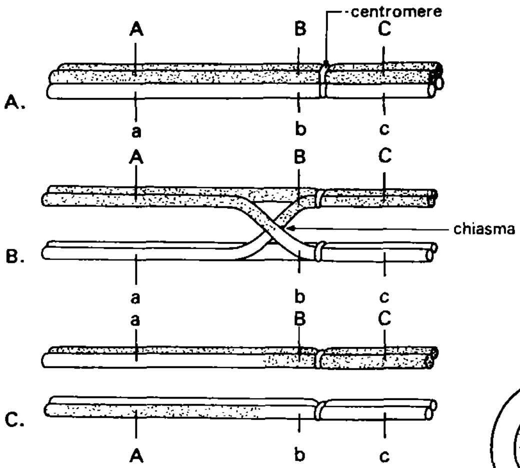
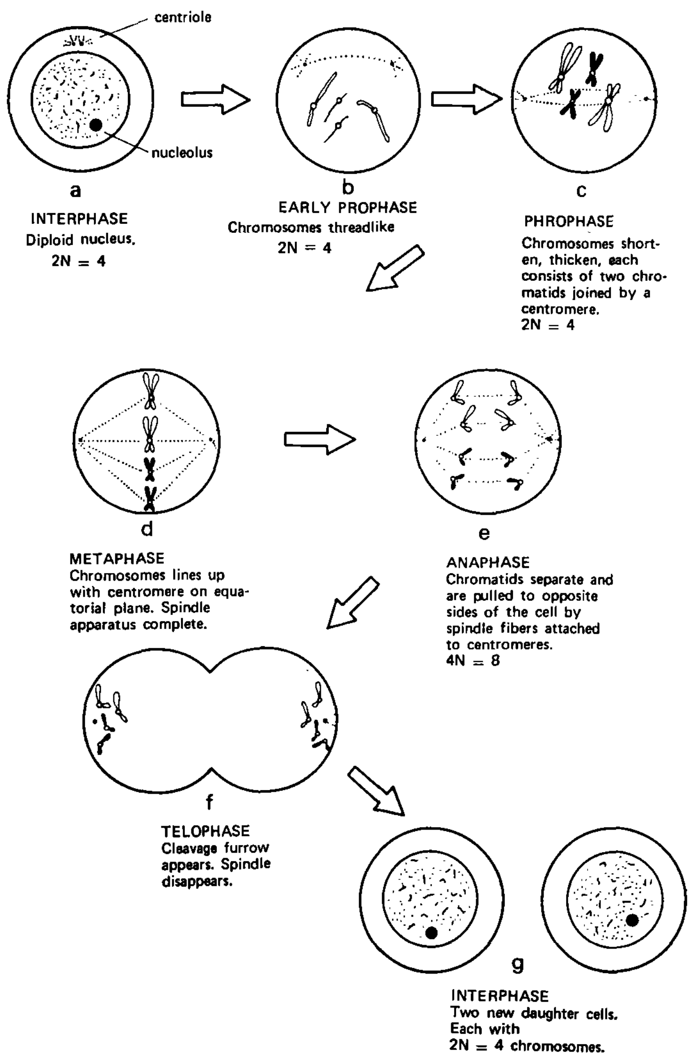


Fig. 17
A. Homologous chromosomes in synapsis forming a tetrad consisting of four chromatids.
B. Chiasma formation as chromosomes separate following crossing over in meiosis.
C. Chromosomes now completely separated. Genes *A* and *a* are now on different chromosomes

Source: redrawn from Paul Amos Moody. *Genetics of Man*. New York, W. W. Norton & Company, 1967.

Fig. 18
Stages of mitosis



chromosome are separated into the final gametes. Thus, each gamete has only half the normal number of chromosomes and each original pair of chromosomes has been separated. Since, for some of the loci, the cell was heterozygous (as *MN* in the example above), the gametes will belong to two types: half will contain *M* and half will contain *N* genes. The gametes of each of the parents, if they were both heterozygous, will be one half *M* and one half *N*. Since the fertilization of an egg by a sperm can occur in any of four different ways, as shown in Fig. 19, and since the two kinds of heterozygous offspring are indistinguishable, three different offspring types may result. Further, since the different types of fertilization are equally probable, the offspring would occur in the following proportions: one *MM* to two *MN* to one *NN*. This, of course, is the offspring ratio observed by Mendel. In a similar way, we may see that all the offspring of parents, both of whom are homozygous for the *M* gene, will also be homozygous for that gene, while if both the parents are homozygous for *N*, their children will also be homozygous for *N*. If one of the parents is heterozygous, *MN*, and the other parent is homozygous for either *M* or *N*, then half their offspring will be heterozygous and the other half will be homozygous for the same gene as the homozygous parent.

Fig. 19
Diagram illustrating the combination of gametes arising in heterozygous parents

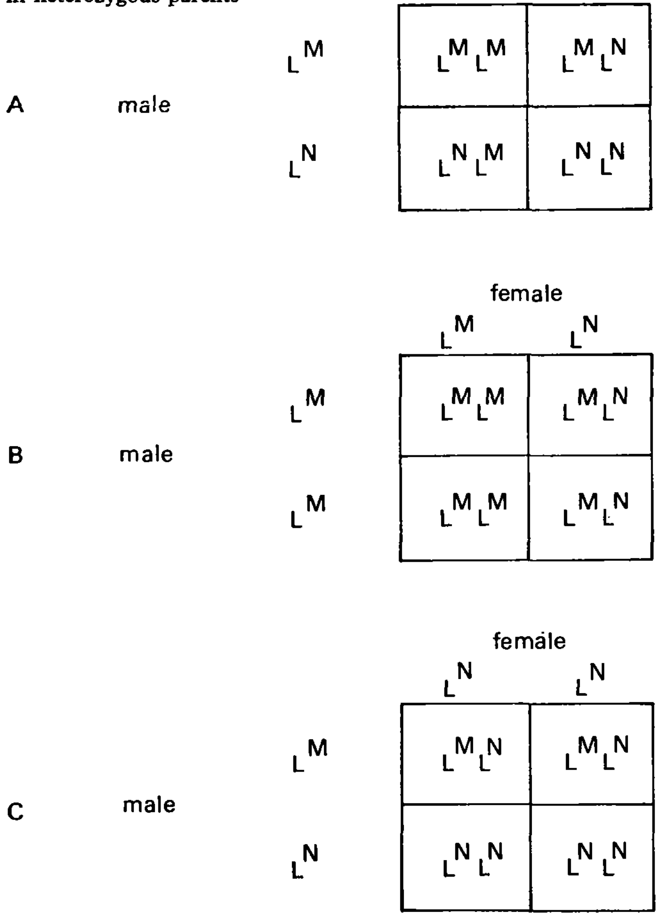
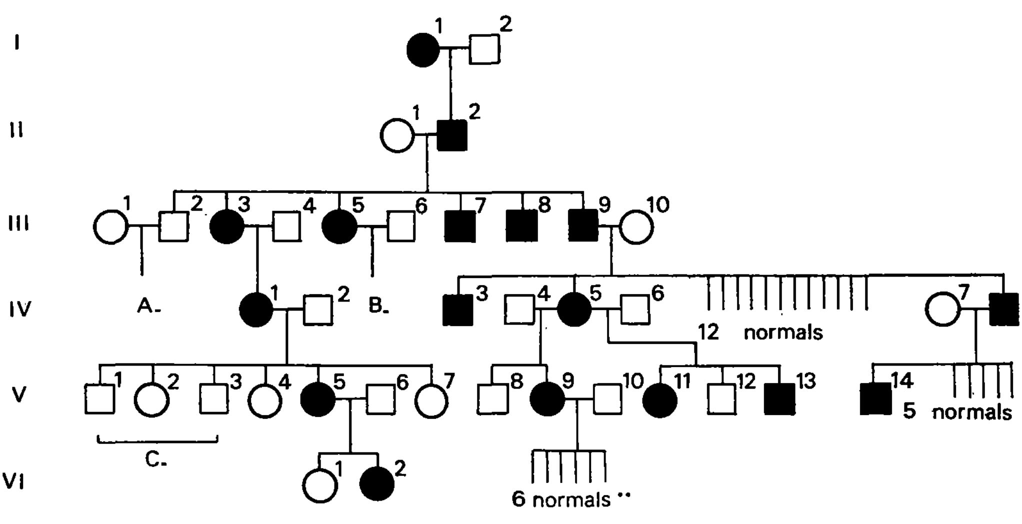


Fig. 20
Pedigree of Huntington's chorea. A. This couple had 9 children and 53 grandchildren, all normal. B. This couple had 11 children, 3 with Huntington's chorea. These three married but did not pass on the trait. C. These three married, and produced a total of 19 children, all normal. (** It is possible that these individuals were young at the time the data were collected. Symptoms may have developed in later life.)
Source: redrawn from Paul Amos Mood *Genetics of Man*. New York, W. W. Norto & Company, 1967.

But it should be recalled that these relative frequencies will be realized only in large numbers of offspring, since they are dependent on chance encounter of the gametes. Thus, they are only probabilities. However, these probabilities are of interest in genetic counselling, since they can be used in decisions on whether or not to have children, and in the study of evolution, where large numbers of offspring are considered.

The kinds of parents and offspring observed in the blood group example above are only recognizable in the case of genes on autosomes and only where the heterozygotes are distinguishable from the homozygotes. The genotype is the actual make-up of genes in the cells or, alternately, the symbols for this make-up, as *MM*. The phenotype, on the other hand, is the visible or measurable expression of these genes in the individual, as the presence of *M*-type antigens in the blood of the individual. In some characteristics, as in the *MN* blood groups, one can readily distinguish what type of genes are present by observing the blood type, but for some characteristics, it is not this simple. A gene may be of a type that is always expressed, regardless of the presence of a different allele at its paired locus. This is a case of complete dominance. That is, the genotype *AA* is not distinguishable from the genotype *Aa* on the basis of phenotypic examination. Such is apparently the case for the *ABO* blood group system in man, where the genes for type *A* and type *B* are dominant to the gene for *O* type. However, since dominant genes for severe disorders are rare (they tend to be eliminated when they are expressed), homozygotes which could result from the matings of two heterozygotes, are often unknown. Dominant genes, such as that for Huntington's chorea, are found only in the heterozygous state in man. But other tests, based on family pedigrees, can be used to determine degree of dominance. In the genetic disorder called Huntington's chorea, the bearer of the gene suffers mental deterioration and involuntary muscular movements. The disease is apparently always fatal, but it usually appears in middle age so that bearers may marry and have offspring. In this case, half of their offspring (if the spouse is homozygous for the normal allele) will also be a bearer of the Huntington's chorea gene (Fig. 20). Thus, if there is a history



of the disorder in one's family, it is possible to calculate the probabilities that the person in question carries the gene for that disease and, if so, that one's child will carry it. If, for example, it is known that one's father has died of Huntington's chorea, the probability is 0.5 that one is a carrier of the gene and 0.25 that a given child of his marriage will carry the gene. This is true, when a gene is completely penetrant, i.e. it is expressed every time it is present.

However, in some characteristics, the controlling gene may meet most of the criteria for a dominant gene, yet occasionally fail to be expressed. If one has ruled out the possibilities of illegitimacy or adoption and if the gene fails expression too frequently to be accounted for by mutation, we say the failure is due to incomplete or partial penetrance. If the gene is found, through studies of pedigrees, to be expressed in 90 per cent of the cases where it is present (as in retinoblastoma, a type of malignant tumour in the eyes), the probability that a given child of a heterozygote will express the disease is $0.90 \times 0.50 = 0.45$.

In autosomal recessive inheritance, the gene responsible is expressed only in the homozygous state, as is the case for albinism. Since the gene may be carried for several generations in the heterozygous state (therefore not expressed), the offspring of parents who are not aware of their heterozygosity for this gene may come as a surprise to the parents. However, assuming that it is known that albinism occurred in the families of both parents and that neither parent is an albino, one can calculate the probability of a child *C* being albino. Thus, if one parent (*A*) has an albino uncle, it may be assumed that the grandparents on that side of the family were both heterozygotes. Then the probability that *A* was a heterozygote is 0.67. (Two-thirds of the non-albino children of two heterozygotes are expected to be heterozygous themselves.) If the spouse of *A* is a normal homozygote, the probability that *A*'s child will be heterozygous would be the probability for *A* to be heterozygous multiplied by the probability that, if he was, his child would be heterozygous, or $0.67 \times 0.5 = 0.34$. Similarly, one calculates the probability that the wife of *A* is heterozygous. Then, the probability that the child is heterozygous is the product of the probabilities of *A* and his wife being heterozygous times 0.5, since half of the offspring of this cross would be expected to be heterozygous. If the probability that *B* is heterozygous is also 0.34, the probability a *C* has of being a heterozygote is $0.34 \times 0.34 \times 0.5 = 0.06$. Similarly, the probability of a given child will be an albino is $0.34 \times 0.34 \times 0.25 = 0.03$.

A further complication arises if the gene is located on the *X*-chromosome. This is the case, for example, of the recessive gene of hemophilia or 'bleeder's disease', in which the blood does not clot normally. Females are rarely affected, since they would have to be homozygous in order to express the disease. This occurs barring mutation only when an affected male happens to marry a heterozygous or an affected female. The male, on the other hand, having just one *X* chromosome, needs only to receive the gene from his mother to be affected. If the male does not have the disease he is not a carrier (heterozygous). One can see, then, if the female in a particular cross has the disease, all of her gametes will carry the gene. Then all of her sons, in this cross, will have the gene (and express it) and all of her daughters will be carriers, i.e. will be heterozygous. If the female is heterozygous, then half her gametes will carry the gene.

Thus, if she mates with a normal male half her daughters will be heterozygous for the gene, none of her daughters will have the disease, and half of her sons will be expected to have the disease. It should also be clear from this that the only way a female could normally be a bleeder would be if her father had the disease and her mother was homozygous or heterozygous for it.

There are also cases where a disease has no clear Mendelian basis, i.e. it may result from the interaction of several loci, some of which may involve multiple alleles. In cases such as these, one does not calculate probabilities, but rather bases forecasts on an analysis of the frequency of expression of the disease in family histories or in the general population. The risks, then, are determined from empirical observation of the disease itself. Such empirical risks are used in genetic counselling in the case of a number of disorders, such as spina bifida, club-foot and Down's syndrome.

From DNA to phenotype

A biochemical chain of events has been found to be effective in producing the trait according to the sequences of units along the DNA molecule. Each three units, for instance, *C*, *A*, *A*, corresponds to an amino acid which will be inserted in the right position in the protein molecule coded by the gene. The sequences of these pairs form the genetic code and serve as templates on which messenger RNA is built (Fig. 21). While the DNA remains intact in the nucleus, messenger RNA, acting as an intermediary, leaves the nucleus, connects itself to ribosomes in the cytoplasm, and serves as a template for the aggregation of amino acids to form protein molecules. In the cytoplasm, amino acid units are connected to molecules of transfer RNA, which attach to their appropriate messenger RNA sites. Of course, during this process, they are aligning their amino acid units into definite sequences. These long chains of amino acids are then released as the proteins.

The proteins are the substances which control metabolic processes (as enzymes) or serve as building materials for cellular structure. The nature of the proteins, then, determines the nature of the cells and, consequently, the physiological and structural characteristics of the individuals. It may be seen that a change in the DNA information results in a change in the messenger RNA, and, consequently, in the amino acid sequence, thus a difference in the resulting protein. The difference in the protein produces a difference in the physiological or structural parameters of the individual.

Mutations are changes in genetic information. One form of mutation is the change in a single gene (gene mutation). Normally, the replication of the DNA molecule involves the unwinding and fission of the molecule between the linking elements, adenine to thymine and guanine to cytosine. The rebuilding of each strand (half) of the original molecule into a complete molecule is accomplished by the addition of an adenine component to each exposed thymine or the addition of a guanine to each exposed cytosine. If a change occurs in any of these linking components, such that, say, an altered guanine is attached temporarily to thymine, the replication of the new molecule will result in a new sequence. Thus, the gene will be altered and will instruct

SECOND LETTER

FIRST LETTER	U		C		A		G		THIRD LETTER
U	UUU	Phenylalanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U
	UUC	Phenylalanine	UCC	Serine	UAC	Tyrosine	UGC	Cysteine	C
	UUA	Leucine	UCA	Serine	UAA		UGA		A
	UUG	Leucine	UCG	Serine	UAG		UGG	Tryptophan	G
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U
	CUC	Leucine	CCC	Proline	CAC	Histidine	CGC	Arginine	C
	CUA	Leucine	CCA	Proline	CAA	Glutamine	CGA	Arginine	A
	CUG	Leucine	CCG	Proline	CAG	Glutamine	CGG	Arginine	G
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine	U
	AUC	Isoleucine	ACC	Threonine	AAC	Asparagine	AGC	Serine	C
	AUA	Isoleucine	ACA	Threonine	AAA	Lysine	AGA	Arginine	A
	AUG	Methionine	ACG	Threonine	AAG	Lysine	AGG	Arginine	G
G	GUU	Valine	GCU	Alanine	GAU	Asparagine	GGU	Glycine	U
	GUC	Valine	GCC	Alanine	GAC	Asparagine	GGC	Glycine	C
	GUA	Valine	GCA	Alanine	GAA	Glutamic acid	GGA	Glycine	A
	GUG	Valine	GCG	Alanine	GAG	Glutamic acid	GGG	Glycine	G

Fig. 21
The genetic code consisting of 64 RNA triplets and their corresponding amino acids. The three triplets UAA, UAG, and UGA serve as 'periods' or 'punctuation marks' in the coded sentences. Note that since 61 words code 20 amino acids, there are 'synonyms'. These synonyms differ only in the third nucleotide

the cell to synthesize a slightly different protein (Fig. 22). The normal process of true replication is an extremely stable one. If this were not true, life could not exist as we know it. Although the changes discussed above are infrequent, they occur at a fairly predictable rate, which depends on the particular kind of mutation.

In one million births, about fifty babies will carry a new dominant gene for retinoblastoma. Since the gene is dominant, it will be expressed in the heterozygote. The frequency of the mutation, then, may be estimated from the frequency of the appearance of the disease. Fifty births out of one million is equivalent to 1 per 20,000. To produce 20,000 births requires 40,000 gametes. Hence, about 1 in 40,000 or 2.5 in

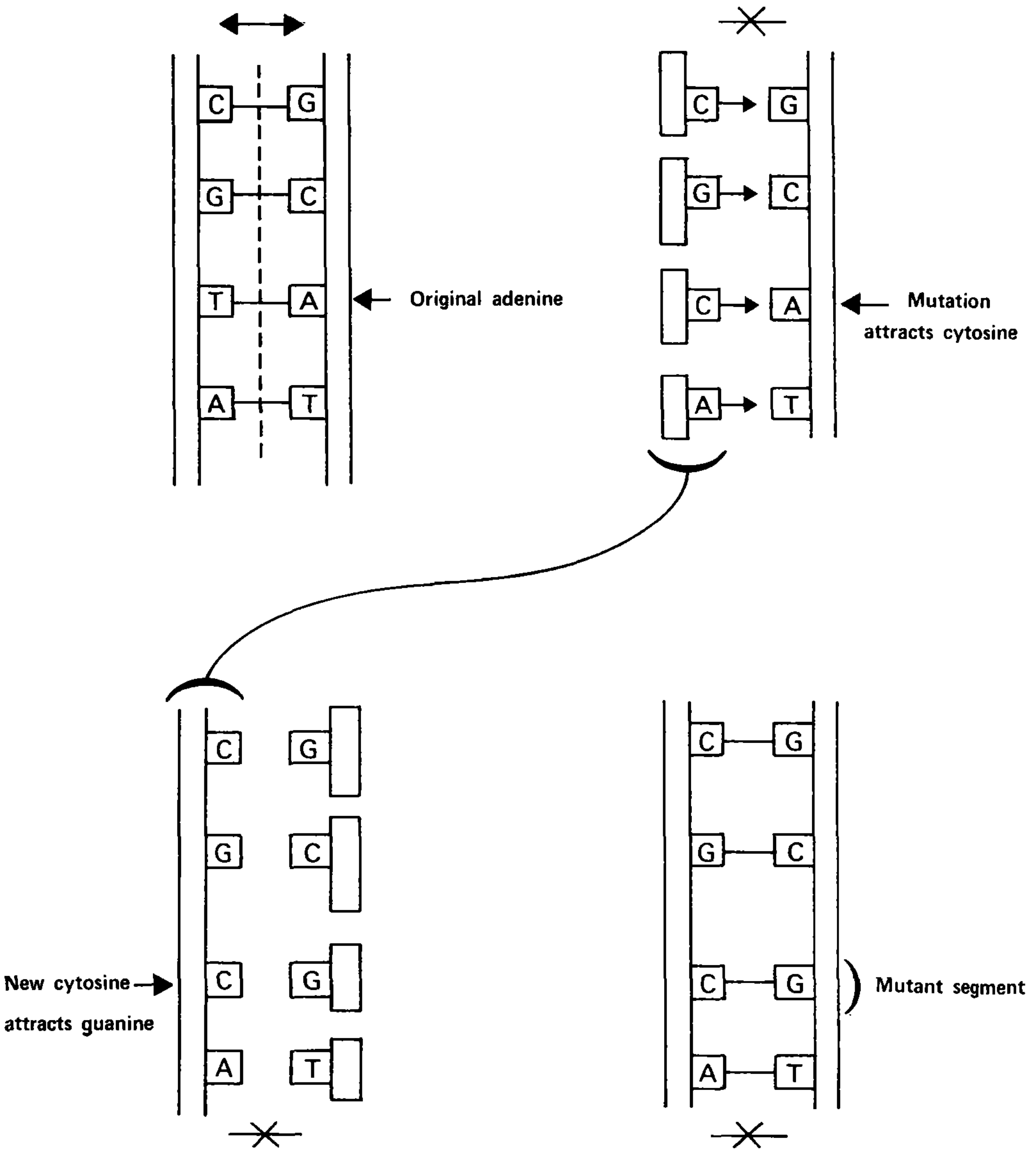


Fig. 22
Model of a genetic mutation through miscopy of the DNA

100,000 carries the defective gene. Since the gene is dominant and leads to death before the patient reaches reproductive age (except when treated surgically) every new affected person used to be originated by a new mutation. Thus 2.5 in 100,000 is an estimate of the mutation rate for this gene.

Achondroplasia is a form of human dwarfism, in which the limbs are abnormally short (Plate 11, opposite page 49). It is caused, again, by a dominant gene, so it is easy to identify the mutation as it occurs. Most of the babies born with the disease die in infancy. Achondroplasia occurs in about 1 child out of 12,000 births. In a manner similar to that used above, it may be calculated that mutations leading to this phenotype occur in about 4 out of 100,000 gametes.

Although the calculations are slightly different for hemophilia, it has been estimated that the mutation (a recessive, *X*-linked one) occurs in about 3 out of 100,000 gametes bearing the *X* chromosome.

These mutation rates indicate that new mutations can affect only negligeably the risk of a child receiving a harmful gene. Thus, these frequencies are usually of more interest in treatments of population phenomena or evolution than in genetic counselling. Other forms of mutation stem from changes in chromosomes due to errors in cell division. They consist of the loss or gain of whole sets of chromosomes, or one or whole chromosomes, or parts of chromosomes (including inversions and translocations of parts). These mutations lead to a considerable amount of change in the genetic information. One of the comparatively common and better-known forms is non-disjunction, the failure of paired chromosomes of sister chromatids to separate and move to opposite poles of the cell prior to division (Fig. 23). Non-disjunction results in daughter cells having both or neither of the affected chromosomes or chromatids. If this process happens in a meiotic division, it can lead to a gamete having too much or too little information (the latter is the more harmful). When a gamete of this kind is fertilized by or fertilizes a normal gamete, the zygote will have either three chromosomes of that type or only one, whereas the normal complement, of course, is two. Other forms of aberration may also result in the inclusion of too much genetic information in a cell, e.g. translocation of a part of a chromosome, which may then be carried through subsequent cell divisions, providing an extra set of genes for that part of the chromosome.

One of the more common human abnormalities resulting from such chromosomal aberrations is Down's syndrome, which results from trisomy of chromosome 21. It has been shown that the syndrome may be caused by simple trisomy (about 95 per cent of the cases) and by translocation (about 5 per cent of the cases). In this abnormality, many characteristics are affected, including facial features, mental retardation, always present, often accompanied by skeletal and heart malformations, and muscular and neural abnormalities. The expectation of life is reduced, although medical treatment has resulted in a steady increase in longevity. In London, in 1932, mean survival was nine years. By 1949, this had reached 12 years. In 1963, in Australia, a study showed this had risen to 18. It is probably higher now. Over a variety of populations, which show remarkably close agreement, the incidence of Down's syndrome is about 1.5/1,000 births. Family studies indicate that the syndrome usually appears as a

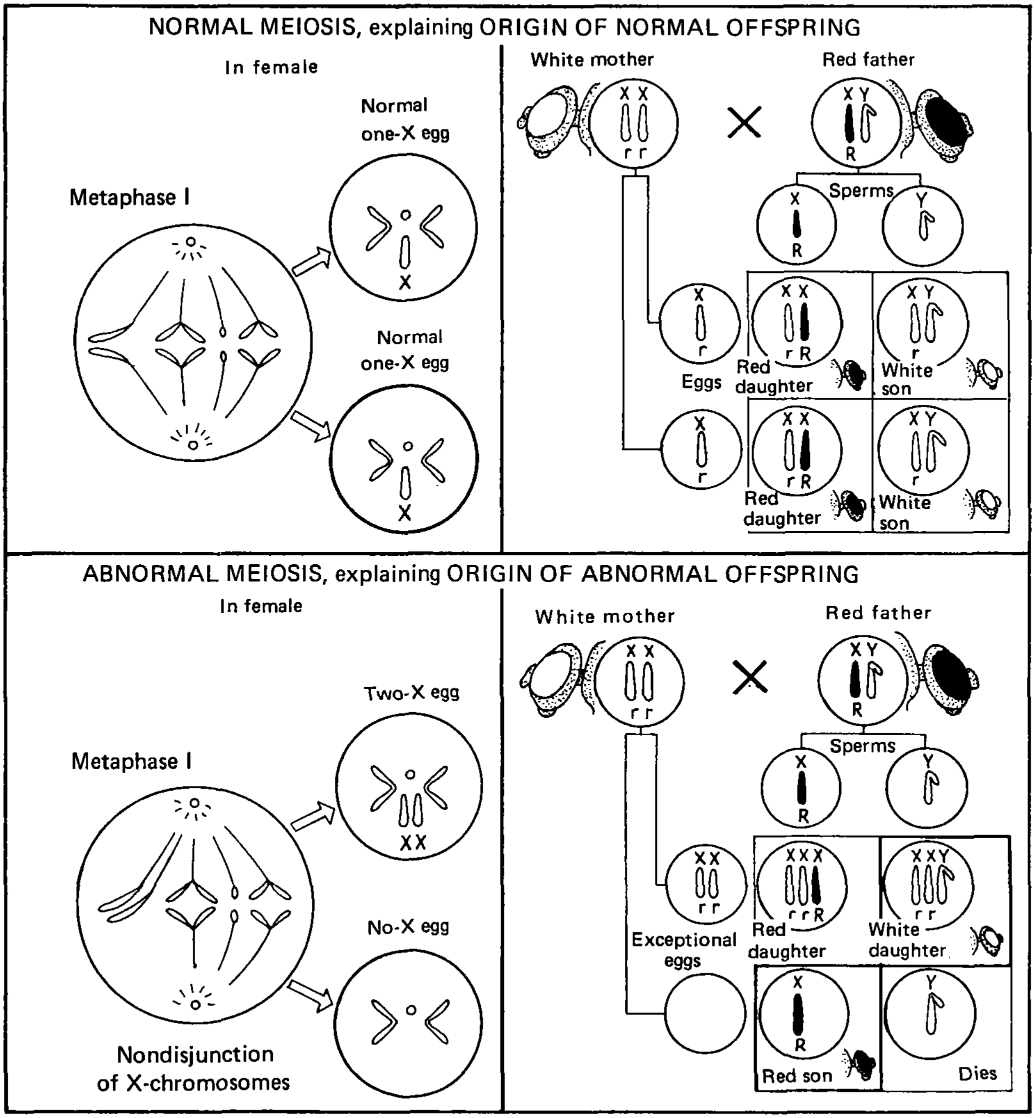


Fig. 23
Non-disjunction of the *X*-chromosome in *D. melanogaster*

Source: G. G. Simpson; C. S. Pittendrigh; L. H. Tiffany. *Life, and Introduction to Biology*. London, Routledge and Kegan Paul, 1959, p. 298 (modified).

Note. The upper left figure shows the normal course of meiosis I leading to eggs that have one *X*-chromosome each. The upper right figure summarizes the normal genetical results of a cross between a white female and a red male. The usual form of the genetic checkerboard has been modified to include a summary of the chromosome, as well as the gene, content of each fertilized egg. The lower left figure shows the abnormal form of meiosis I, which Bridges in his work on *Drosophila* guessed was the cause of the exceptional offspring he had found. On the right the genetical results *predicted* by this hypothesis are given in the form of a genetic checkerboard. The two squares in heavy outline explain the exceptional offspring and lead to the crucial prediction which made this study of Bridges a landmark in genetics. What is the prediction? For diagrammatic simplicity we have represented the oöcyte (at metaphase I) as giving rise to two eggs. In reality, an oöcyte gives rise to only one egg. This does not, however, affect the implications of the diagram. Thus, while in the lower half of the figure, we show the abnormal oöcyte giving rise to *both* a two-*X* egg and a no-*X* egg, in nature it gives rise to *either* one or the other. But since a female lays large numbers of eggs she will produce in the long run *both* the two-*X* and the no-*X* eggs as the result of nondisjunction.

sporadic event and since the individuals usually do not mate (there are recorded exceptions), the abnormality is not transmitted. In about 5 per cent of the cases, however, the risk of recurrence increases because the trisomy results from a translocation carried in balanced state in a normal member of the family. For simple trisomy, the probability of a second child showing the disease if an earlier child is affected is approximately the same as the frequency in the population, except for age effect. The incidence of Down's syndrome produced by simple trisomy increases with the age of the mother, independent of other factors. In mothers up to 30 years of age, the incidence is about 0.7/1,000 births and about 3/1,000 births over age 30. If only mothers 45 years old or older are considered, the frequency of occurrence is about 1/50. This information is useful, of course, for counselling mothers on the empirical risk of a second affected child where the first one showed simple trisomy-21.

The determination of empirical risk in the cases of Down's syndrome being produced by translocation is complicated by the fact that there are several different types of translocations which can result in trisomy-21. In these cases, the genetic counsellor would acquaint himself with the type shown in the first affected child, and any abnormality in the chromosomes of the parents, then counsel the parents on the associated empirical risks accordingly.

Genetics counselling

When a baby is born with a defect, or when a disease which could be hereditary in nature appears, the family of the affected is concerned not only with its prognosis and treatment, but also with the possibility that the defect or disease will turn out in another member of the family. The probability that a particular genetic abnormality will occur can be estimated on the basis of Mendelian family laws or on the results of empirical studies. The purpose of genetic counselling is to make such estimate and discuss it with the people concerned so that those people in the family contemplating having a baby may learn beforehand the risks involved and plan accordingly.

There are, however, certain abnormalities which appear to be of genetic origin, but are not. Intrauterine infections such as syphilis or German measles, may affect embryos but are not genetic in nature. Also, deformities in offspring resulting from use of drugs by pregnant mothers are not genetic disorders. These affect the already forming embryo.

Genetic counselling is the applied side of human genetics. It is an important service given by specialists to people under risk of having affected children. Both by giving them the information which makes them decide not to reproduce or showing that they are overconcerned about a trivial risk, the genetic counsellor helps to alleviate or preclude much unhappiness.

3 Mechanisms of evolution

What is evolution?

In a biological sense, evolution is understood as the gradual change in populations of the frequencies of genes and corresponding traits, resulting from mutation and recombination of genetic materials under the influences of natural selection and genetic drift. The functional unit on which it operates is the population, which may be defined as reproductively somewhat isolated group of organisms of the same species. Complete isolation is not necessary for differentiation to occur, with only the rate of diversification being affected by the degree of genetic exchange with other populations.

The genes collectively possessed by all the individuals in the population at a given time may be regarded as a 'gene pool' from which the gametic sample will be drawn in order to produce the next generation. Reduced to the simplest treatment, the pool consists of the genes possessed by the population for a particular locus, each individual being represented by two genes, since he is diploid. If there are only two possible alleles for that locus, there are three types of individuals in the population. We may identify the two different alleles as A and a and the three diploid types as AA, Aa and aa. The group of genes will then consist of some part a genes and the remainder will be A genes. The relative frequencies of the two types will be determined by the relative frequencies of the individuals of types AA, Aa and aa in the population. Note that the proportions of A and a may be the same even though there are different proportions of AA, Aa and aa individuals (Fig. 24), since the frequency of A is the sum of the frequency of AA and half the frequency of Aa. The frequency of a is similarly determined by the Aa and aa members of the population. Intuitively, we may extend the above considerations to include cases where more than two alleles occur for a particular locus and where many loci are considered simultaneously.

Since the genetic constitution of a population at a given time is derived from the gene pool of previous generations, genetic continuity is maintained through time; however, small changes in the gene pool are constantly induced by a number of factors and these are the sources of shifts in adaptation which permit the population to continue to survive in a changing environment. There is a relationship called the Hardy-Weinberg law which provides some insight into the nature of the mechanisms

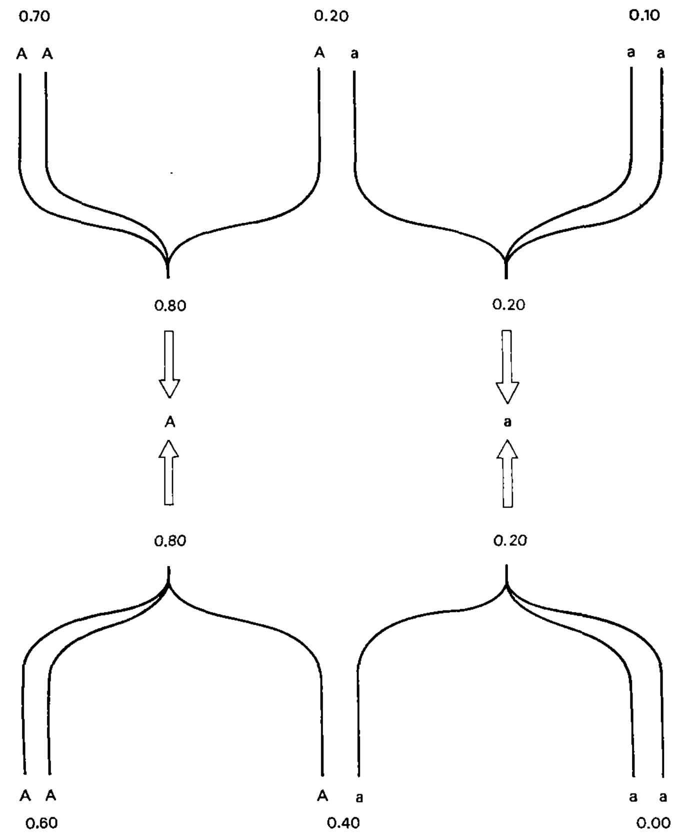


Fig. 24
Gene frequencies of *A* and *a* with two different phenotypic frequency arrays

which bring about these changes. Since a change in the proportion of *A* genes and *a* genes in the population is evolution, we might ask what factors can cause these changes. The Hardy-Weinberg law states that, in a very large random-mating population where there is no mutation, no migration and no selection, there will be no changes in gene frequencies (hence, no evolution). This, then, indicates what factors can cause evolution.

If the population is small, random effects of gamete selection cause genetic drift. The effect of non-random mating is more complex and tied to natural selection, which is the favouring of some kinds of genes by the environmental circumstances. Migration can obviously import to, or export from, the population more *A* than *a* genes or vice versa. And mutation, by its nature, produces an automatic change in gene frequency. One may see, then, that these factors are those that are effective in changing the gene frequencies of populations, or, in other words, bring about evolution.

The many populations of a species are constantly evolving, while the selection imposed by the environment may remove or reinforce one or more of the populations to the benefit or detriment of the remainder. Thus, evolution on the species level results from changes in the gene pools of its populations with the occasional replacing of some populations by those which are more fit. In this manner the differentiation of races and eventually speciation begins.

Mutation and recombination

Mutation is a change in the basic genetic material. It can be a change in a DNA molecule, an alteration of the relationship between genes within a chromosome, a change in the number of chromosomes, or the addition or loss of parts of chromosomes. Mutation represents a disruption in the genetic continuity of a population. Hence, it is an automatic source of 'newness' for the population, or, in other words, a source of evolutionary potentiality, which can reduce or increase the degree to which the population is adapted to its environment. The great majority of these mutations are deleterious to their bearers and tend to be eliminated, but some can contribute to better fitness and be favoured by natural selection.

Another form of originality for the population is the recombining of existing material. For example, if we have two people who are of the genotype *Aa* in a population wherein all the individuals are either *AA* or *Aa*, and if these former two should mate, some of their offspring would probably be of the genotype *aa*. Phenotypically, these offspring will be different than any previous members of the population, or new, even though only previously existing genetic material is used to produce them. Since natural selection operates only through phenotypes to change the make-up of the gene pool, a new rate of change may occur. But it should be clear that the recombination depends on the availability of different types of genetic material, leaving mutation as the ultimate source of variability, hence of biological originality. Any differences among genes are the accumulated results of mutation over the history of the species.

Since mutations are almost always harmful, one could wonder how more than a million species of organisms arose from spontaneous changes in genetic information through about 4,000 million years. We must remember, however, that in any one time period (say, a year), many individuals are being produced. Moreover, the rates we have been discussing are for only one locus. In man, if there are at least 25,000 loci (probably a very conservative estimate) and 2 mutations occur for each 100,000 gametes for each locus, we may expect that a substantial number of our gametes contain one or more new mutations of some sort. Thus, the power of mutations as a factor of evolution gains credibility.

Genetic drift, the chance component of evolution

Let us take a simple Mendelian characteristic which is controlled by a single locus with only two possible alleles, say *A* and *a*. We know that under these conditions, barring unusual circumstances, the gametes produced by the parents will be in the same frequency with regard to *A* and *a* as are the frequencies of *A* and *a* in the gene pool. But, not all the gametes will be used to produce the new generation. Instead, if they are in small numbers, i.e. the population is a small isolate, in the random sample of gametes used for actual fertilization, it is very unlikely that precisely the same frequencies of *A* and *a* will result. Rather, the more likely occurrence is that the

sample will present a slightly different frequency, say 0.68 of *a* if the parental frequencies were 0.70. This change in the make-up of the gene pool from generation due to the sampling of the gametes from the array of possible gametes leads to 'genetic drift'. Note that this change is a result of the random sampling of gametes alone and has no bearing on whether the gene is harmful or beneficial. In Fig. 25, we see that the change occurs in each generation independently of the direction of change in the previous generation. Hence, one cannot know the ultimate outcome of the changes. In the event that the sequence proceeded as the solid line in Fig. 25, the frequency reaches 1.0, the gene is 'fixed' and its allele is lost to the population. (It could as well have occurred the other way around, with the frequency for *A* at 0.0.) The lost allele then may only re-occur by mutation or by being re-introduced from another population.

This is in agreement with the law of 'central tendency', which works in the following way. Suppose we had a container of seeds, half of which were white and half of which were red (and which differed in no other way); we close our eyes, mix them

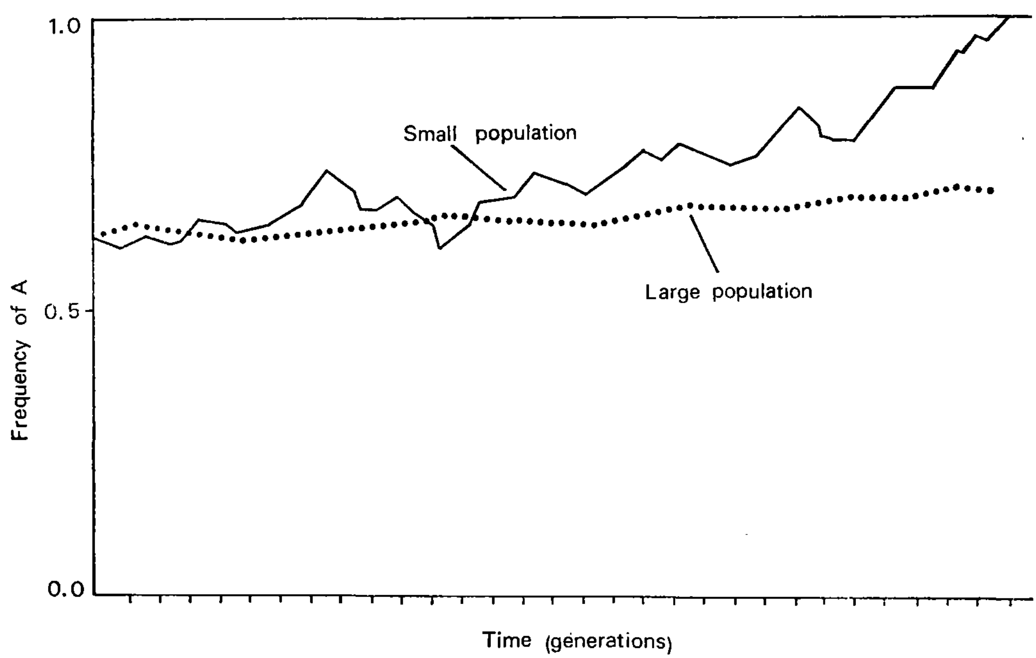


Fig. 25
The change in gene frequencies with time as a consequence of genetic drift

up and take a handful. Upon opening our eyes, we count them and determine in our sample the proportion of, say, the red seeds. We would find that they would not be exactly half red. In fact, they might be only, say, 42 per cent red seeds. Now we put the seeds back into the container, mix again and take another handful with our eyes closed. We get another figure. If we repeat this many times, two things will become

obvious. First, each handful would be somewhere close to 50 per cent red, and, secondly, if we calculated the average of successive tries, we would likely find that the average would come closer and closer to 50 per cent. In other words, the larger our sample size the closer we are likely to come to the expected value. The mechanism of genetic drift works in precisely the same way and for the same reasons. In a small population, as that depicted by the solid line in Fig. 25, the sample of gametes is small in each generation. Hence, the outcomes are quite variable. In larger populations, as that depicted by the broken line in Fig. 25, the changes from generation to generation are quite small. Thus the effectiveness of genetic drift is dependent on the population size. In a very large population, the effects are negligible. But one should bear in mind that even in a very large group, there may be smaller sub-populations that each constitute an isolate or mating group. In this case, whereas the frequency of a gene is drifting down in one group, it is likely to be independently drifting up in some other group, so the over-all effect is that for the large group, there is little change, even though there may be a considerable differentiation among the sub-populations.

Natural selection

Although the treatment of underlying mechanisms and rates of operation have been refined in this century, the meaning of natural selection is essentially the same as that which Darwin proposed, i.e. the favouring or disfavouring of certain traits by the environment. This environment is taken to include other organisms as well as the physical and chemical surroundings. It is perhaps most clearly exemplified in the case of hereditary diseases. If a particular genetic make-up results in a child being born with deformed legs, and if that malformation prevents the child from competing on an equal basis with other members of the population for food, or in some other way reduces the probability of that child surviving to reproduce, then the genes responsible will tend to be removed from the population. The sole criterion for selection is whether or not the individual involved leaves offspring which can survive to reproduce. The mechanism, then, operates on the reproductive capacity of the organism. Any genetic composition which tends to reduce the probability that its possessors will reproduce will be selected against. Conversely, a genetic composition which increases the likelihood of leaving offspring will be favoured by natural selection. This relative likelihood of participating in reproduction is called fitness or adaptive value. The concept of fitness is usually applied to an individual genotype as compared with the alternative genotypes with regard to a particular locus, all other loci being equal. Then fitness can be defined as the ratio of the number of offspring left by a genotype to the average number of offspring left by all genotypes. For example, if the genotype *AA* is taken as the base and leaves the normal number of offspring, say, 3, its fitness is unity. If the heterozygote (*Aa*) is somewhat less fit, it may leave only 2 offspring on the average, so its fitness is 2/3 or 0.67. If the remaining genotype, *aa*, leaves only one offspring, its fitness is 1/3 or 0.33. Using this means of expressing fitness, we find that it may take any value from 0.0 (complete lethality) to 1.0 (normal fitness) or even

greater (1.1, etc.) for individuals of exceptional fitness. The relative fitness is the complement of the selection coefficient s which measures the intensity of selection pressure against a particular genotype, i.e. fitness is equal to $1-s$. When selection is zero (normal phenotype) fitness is 1 and when selection is 1 (lethality) fitness is zero. The rate of natural selection depends on the type of genetic system with which we are concerned. Mathematically, the simplest case is one of complete dominance, in which the *aa* genotype is lethal ($s=1.00$) and in which the heterozygote (*Aa*) has the same fitness as the homozygous dominant (*AA*). Then, for *AA* and *Aa*, $s=0.00$. This is illustrated by the case for infantile amaurotic idiocy or of Duchenne muscular dystrophy, wherein the affected child always dies before reaching reproductive age. On the other hand, there is not a known selection force against the other genotypes (*AA* and *Aa*).

When q, the proportion of the recessive allele a, is large, the reduction in q per generation is large, but as q becomes small, the rate slows down. The special case cited above leads to a simplification of calculations, wherein n, the number of generations required for q to be reduced by half, is equal to $1/q_0$, where q_0 is any initial value of q. Thus, in the case of infantile amaurotic idiocy, if it did follow these conditions exactly, and if we started at a frequency of, say, $q=0.007$,

$$n=\frac{1}{0.007}=143 \text{ generations.}$$

Thus, it would require 143 generations to reduce the frequency of a from 0.007 to 0.0035. If the mean generation time for humans was, say, 30 years, then this reduction would take about 4,290 years.

If there was partial selection against the heterozygotes, this rate of reduction would be increased. On the other hand, if the intensity of the selection against *aa* were less than unity, the rate of reduction in a would be slower. The latter is usually the case for slightly detrimental genes, for which the value of s is often of the order of 0.01 or less. In this case, the removal of the a genes from the gene pool would take millions of years. Because genes of this type are numerous, their combined effect is substantial, but not as apparent as for lethal genes.

However, since environments are not stable, the intensity of natural selection changes with time and in different fashion in each place. So it is unlikely that the rate of removal of a detrimental gene would remain constant. This does not, however, reduce the credibility of the statement that natural selection changes gene frequencies and, since populations consist of finite numbers of individuals, it can eliminate genes.

Consanguinity increases the likelihood of homozygosity in the offspring of deleterious recessive alleles. This has the effect of increasing the frequency of people affected by recessive ailments and exposing the corresponding alleles more fully to elimination by natural selection. On the other hand, any factor which makes the rate of consanguineous marriages become smaller tends to reduce the incidence of all types of recessive defects and diseases. Two of these factors typical of modern times are the reduction in average family size and the breakdown of isolates, i.e. the mixing of populations which were previously isolated. Since both are in action practically all

over the world and they influence the whole gamut of recessive detrimental traits, the benefit they do to humanity is of considerable magnitude. Concomitantly, however, the recessive alleles become less exposed to homozygosity, hence to selection and this leads to their accumulation in the heterozygous form in the population, increasing thus the potential genetic load of mankind. This increase will take place over the millennia until a new equilibrium is reached where the excess of harmful recessive genes will be manifested in homozygosis and be eliminated at the same rate they are produced by mutation. On this remote occasion, the frequency of affected people will be the same as now, if no new increase in consanguinity happens. In the meantime, breakdown of isolates and decrease in family size have acted as powerful eugenic factors, surpassing by far any intentional measures that could be adopted against harmful genes. It also becomes clear that genetic counselling, while of paramount importance to the family involved, has no pretence to be a significant eugenic factor at the population level. Natural selection, as slow as it is, is much more effective. Another important trend in modern times is the relaxation of natural selection by current medical practices. With the spread of effective medicine comes the retention of more and more alleles which would otherwise have been held in check by natural selection. The over-all effect is to increase the number of individuals exhibiting hereditary disorders, while decreasing, through medical care, the individual severity of each one. The real change in terms of human stress and suffering, is that inconvenience for a greater number of individuals will take the place of the suffering of a smaller number of them. This benefit continues provided the medical practice continues as it is. If, however, the practice was discontinued, the human suffering, until readjustment through natural selection, would be enormous. More subtle but perhaps not less effective than this influence of medicine is the repercussion of modern technology in the genetic make-up of the human species. All comfort it brings in terms of defence against the environment and facilitation of the tasks needed for survival will create a corresponding relaxation of natural selection with the accumulation of genes less fit for life in primitive conditions. Again, as long as technology keeps going no harm is done but if some day it suffers a breakdown, humanity will be biologically less fit to return to the jungle and the caves.

Balanced polymorphism

If a gene is at all harmful, we would expect it to be reduced in frequency until it was maintained at a very low equilibrium level supported only by its mutation rate. This might not happen in a small population due to genetic drift. But there is another factor which keeps certain genes at an apparently abnormally high frequency in spite of their being harmful. This is because, in some cases, the heterozygote is more fit than either of the homozygotes. For example, in many African groups, the gene Hb^S is maintained at a frequency out of proportion to what could be sustained by mutation alone. It is responsible for the production of an abnormal form of hemoglobin, type S, whereas the normal hemoglobin is of type A. People who are homozygous for the

gene Hb^S have only type S hemoglobin and display a characteristic deformation of the red blood cells to a sickle-like shape under conditions of low oxygen tension. After sickling the red cell is destroyed, leading to anemia. In addition, S hemoglobin has a reduced capacity to carry oxygen. Those people homozygous for the sickling gene probably do not normally live to reproductive age, but with modern medical treatment some of them do reach adulthood. People who are heterozygous ($Hb^A\ Hb^S$) have both types of hemoglobin. When the red blood cells of these people are subjected to severe oxygen deficiency *in vitro*, they also show sickling. In spite of these apparently harmful effects, as many as 40 per cent of the people in some populations are heterozygotes ($Hb^A\ Hb^S$). The high frequencies of the Hb^S gene are explained at least in part by the protection afforded the heterozygote against *P. falciparum* malaria. In the 'normal' homozygotes ($Hb^A\ Hb^A$), this malaria produces a higher rate of infant mortality than in the heterozygote ($Hb^A\ Hb^S$). Although the rate of infection is about equal in all genotypes, it seems that during the first three years, the parasite count is much lower in the heterozygote than in the normal. Because of this, it appears that the more serious complications of malaria (e.g. cerebral malaria) occur less frequently in the heterozygotes. Therefore, the difference lies in the fact that *P. falciparum* malaria is not usually fatal in the heterozygotes. Thus, those who are homozygous for Hb^A often die of malaria, those homozygous for Hb^S usually die of sickle-cell anemia, and the heterozygotes are favoured, since they are less subject to either problem. The life style of some of these populations also contributes to maintaining the high frequencies of Hb^S in their gene pools. The practice of agriculture promotes an increase in the populations of mosquitoes, which serve as vectors for malaria, leading to increased rates of malarial infections. This intensifies the natural selection effects of malaria, thus serving to maintain a high frequency of Hb^S.

Fluctuations in population size

In non-human populations (and likely in pre-human and early human populations), another factor modifies the rate of removal of deleterious genes. It is dependent on the usual fluctuation in numbers that is characteristic of animal populations.

During those times when the population is reduced to very small numbers, the degree of inbreeding is comparatively high. This results in a much greater phenotypic expression of any recessive genes that occur with low frequency. Since most harmful genes fall in this category, they will be exposed to natural selection in much higher proportions during this time. Thus, when the population is reduced in size, harmful genes tend to be removed rapidly by natural selection.

During periods of population expansion, on the other hand, the rate of manifestation of the remaining harmful genes becomes smaller due to the reduction of inbreeding. When the population achieves maximum numbers, very few members will show the effects of these genes, so survival rates are high. However, new mutant recessive genes will accumulate through this period of low inbreeding. Thus, the periods of high numbers offer short-term benefits in the sense of genetics, but in the

long term, contribute to the accumulation of deleterious genes. The periods of low numbers will be characterized by relatively high frequencies of spontaneous abortion and deformed offspring, but offer, in the long run, a better adapted population as a result of reduction of harmful genes. These fluctuations, then, regulate the balance between a limited genetic load and a reasonable variability, which is useful for rapid adaptation to changed environments.

These two components of evolution operate independently, although both change the frequency of genes in the pool and it is difficult to separate their effect empirically. Whereas genetic drift changes gene frequencies without regard to the effects of the gene on the organism, natural selection works on the inherent value of the gene for the adaptation of the organism. The relative effectiveness of both factors is strongly affected by the size of the population. The larger the population size, the more important is selection in relation to drift. Conversely, in very small populations, such as on small islands, genes may be readily fixed or lost with little reference to adaptive value, so that drift is more effective than selection for differentiation among island groups which are otherwise exposed to seemingly very similar environments.

The real rate at which genetic drift occurs in small populations has been studied in the case of the Dunkers, a small religious isolate which migrated to the United States from Germany less than 300 years ago. Their isolation has been maintained until now, through religious practice, which forbids intermarriage with other groups. A comparison of the gene frequencies among the Dunkers with the frequencies of the same genes in the source population and the surrounding ones shows that an appreciable amount of genetic drift has occurred.

Bearing in mind that large 'populations' may consist of small breeding groups, since they almost necessarily occupy large areas, local variations might still be brought about by genetic drift. This phenomenon is also related to the life style of the people involved, as is the case of the Dunkers. That is, any cultural factor which affects the mating pattern could also affect the degree to which adjacent sub-populations exchanged genetic material. Thus, the effective size of the population (or sub-population) is influenced by the life style, which therefore affects the relative magnitude of intensity of genetic drift and natural selection.

In pre-man, the hunting groups were likely limited to small populations, rather isolated from each other by the need for a fairly large supporting area. The early hunting bands probably consisted of groups smaller than that of the Dunkers. Therefore, genetic drift was likely quite effective in the differentiation of these small groups. Natural selection was probably more effective at that time in the favouring of one competitive group over another rather than operating within each group. As man changed his life style from the small hunting band to larger ones, the effectiveness of genetic drift as a component of evolution was reduced. The development of more sedentary cultures based on food gathering rather than hunting and later based on agriculture, increased population size and therefore further reduced the rates of genetic drift, except where some other aspect, such as religious isolation or other inbreeding systems, came into play.

In modern man, except for some unusual island populations, cultural patterns

would be the main existing isolation mechanism. With the improvement of transportation means, increased travel among many people of the world has anihilated, for the time being, any significant influence of genetic drift on the alteration of gene frequencies in human populations.

Changing environments

Since environments are not indefinitely stable, adaptation is temporary, at least in terms of the geological time-scale, whether for man or any other species. In confronting a change in the environment, a species either becomes sufficiently adapted to the new conditions or perishes. The fossil record provides ample evidence that the latter case has been frequent. Indeed, environmental changes bring about evolution, but involve difficulties and risk for a species.

First, mutations, the ultimate source of variability, do not arise because they are needed. That is, the change in the gene is a spontaneous event which is usually harmful. Overcoming a rapid change in the environment, then, by the advent of new mutations is an extremely unlikely event. Secondly, natural selection has tended to remove the previously harmful genes from the species. But genes which produced characteristics ill-adapted to the old conditions may be desirable in the new environment. Sickle-cell anemia, which is produced by a gene that carries a distinct disadvantage for people in most environmental conditions, and is frequently lethal in the homozygous state, provides a marked advantage for people who live under malaria-infested conditions. In addition, the nature, direction and rate of changes in the environment are not predictable with certainty, even by modern man, so that preparing for it by rational artificial selection would be at least impractical, if not impossible.

These and other considerations lead us to the conclusion that the means by which a species survives a major, rather rapid, environmental change is to be prepared, or preadapted, by having sufficient genetic variability on hand so that at least a portion of its members will be able to continue in the new conditions. The individual himself does not evolve to meet new exigencies, but a population may evolve by shifts in the gene pool frequencies through the spreading of the offspring of special types of individuals. The price paid by the species for retention of variability is a certain genetic load manifested by the occasional loss of zygotes and the presence of deformities in some surviving offspring. A large proportion of spontaneous abortions are due to lethal point mutations and to chromosomal aberrations (the latter are estimated to account for 30 to 40 per cent of spontaneous abortions). Of the surviving offspring, some will exhibit mental defects, congenital malformations, or inherited diseases. Still others may be sterile due to genetic defects, as in Turner and Klienefelter syndromes.

By the beginning of the Pliocene epoch, perhaps 12 million years ago, some man-like apes had probably appeared. During this epoch, which lasted about 11 million years, these apes were faced with what might be termed a rigorous environment. The rise of the human species probably occurred at these stages in the African savannahs. During the Pliocene, these savannahs were gradually increasing in extent, the forest

meanwhile receding in the face of a gradual reduction in rainfall. Lack of water would have been a severe problem to the ape-man, as would have the scarcity of food. They had acquired a taste for meat and were dependent on the small animals they could catch. Conditions must have been extremely harsh in those times. But in one respect, the environment was kind. It was quite stable. Although volcanic activity and the formation of the Great Rift Valley would have provided moderate diversion, once a viable pattern of life was established, it could be continued for millions of years. Thus, the stable, dependable environment of the Pliocene would not foster rapid evolution. But during this time, mutations would have been occurring and would have been subjected to selection pressures. Undoubtedly, many were eliminated, although at any given time, some variability would persist. Furthermore, the biological developments which later would permit social development were being formed. With the advent of the Pleistocene epoch, which involved more environmental change, the ape-man had limited, but sufficient genetic variability, which could be expressed through sexual recombination, to adapt to these new climatic changes. In the next million years, he was confronted with four ice ages and their accompanying pluvial conditions, and the need for much more rapid adaptation.

Further reading

BAKER, Paul D.; WEINER, J. S. (ed.). *The biology of human adaptability*. Oxford, Clarendon Press, 1966.

BORDES, F. (ed.). *The origin of Homo sapiens*. Paris, Unesco, 1972, 321 p.

DASMANN, R. F. *Planet in peril?* (Man and the biosphere today). Penguin Books/Unesco, 1972, 135 p.

HANDLER, P. (ed.). *Biology and the future of man*. Oxford University Press, 1970.

LAMBERT, G. *Adaptation physiologique et psychologique de l'homme aux conditions désertiques*. Toulouse, Centre d'hémotypologie du Centre hospitalier universitaire, 1968.

MALHOTRO, M. S. (ed.). *Human adaptability to environments and physical fitness*. Madras, Defence Institute of Physiology and Allied Sciences, 1966.

MEAD, M. (ed.). *Cultural patterns and technical change*. Paris, Unesco, 1953, 348 p.

RASLANSKY, J. D. *Genetics and the future of man*. New York, Appleton Century, 1966.

SANKALIA, H. D. *Some aspects of prehistoric technology in India*. New Delhi, Indian National Science Academy, 1970.

SORRE, Max. *Les fondements biologiques de la géographie humaine. Essai d'une écologie de l'homme*. 4e éd. Paris, Armand Colin, 1971.

VAYDA, A. P. (ed.). *Environment and cultural behavior*. New York, Natural History Press, 1969.

VORSTER, D. J. M. (ed.). *Human biology of environmental change*. London, I.B.P. Central Office (7, Marylebone Road, London, N.W. 1), 1973.

UNESCO. *Traditional cultures in South-East Asia*. Bombay, Orient Longmans, 1958, 270 p.

YOUNG, J. Z. *An introduction to the study of man*. Oxford, Clarendon Press, 1971.

Part II

Environment of human populations

4 Energy

Introduction

Energy is the prime mover of the universe and man tries to know more about and understand this phenomenon in order to control, conserve and use it in everyday affairs. Sources of energy are found chiefly in the radiation from the sun; various types of fossil fuels (oil, coal, natural gas), wood and charcoal; in water, wind, food substances and in the nuclear energy obtained from nuclear reactors. In Asia, water-harnessed energy (hydro-electric) is produced in almost every country, but windmills are not so common, except in some village communities where they are used, to lift water from below ground level, in order to irrigate and serve as domestic water supply. Windmills can also generate electricity when designed to do so.

Human populations, like those of other animals and plants, are intricately involved in the energy relations of the whole earth.

Energy is a fundamental physical concept; the term is derived from the classical Greek, *energeia*, *en* and *ergon*=work; thus, in the scientific connotation, energy is the equivalent of, or the capacity to do, work. Indeed, the physical dimensions of energy, namely ML^2/T^2 [a formula signifying (mass)$\times$(length)^2/(time)^2] are the same as those of work.

Different kinds of energy show themselves in both physical and biological systems; potential and kinetic energy in mechanical processes, surface energy, electrical energy, chemical energy, pressure-volume energy, radiant energy, sound and gravitational energy. The law of conservation of energy is experimentally demonstrable in living systems. Thus, energy may be converted from one form into another, the conversions being equational, i.e. quantitatively equal. The first law of thermodynamics, formulated from these observations, states that energy may be changed from one kind into another, but it cannot be either created or destroyed.

Both the concepts of energy and work are best understood from knowing how to measure them and to show the relationship between them. Units which are used to measure forms of energy are given in Fig. 26.

Efficiency, in the conversion of energy, within the various physical and biological systems, is not very high. For example, an internal combustion engine is generally less

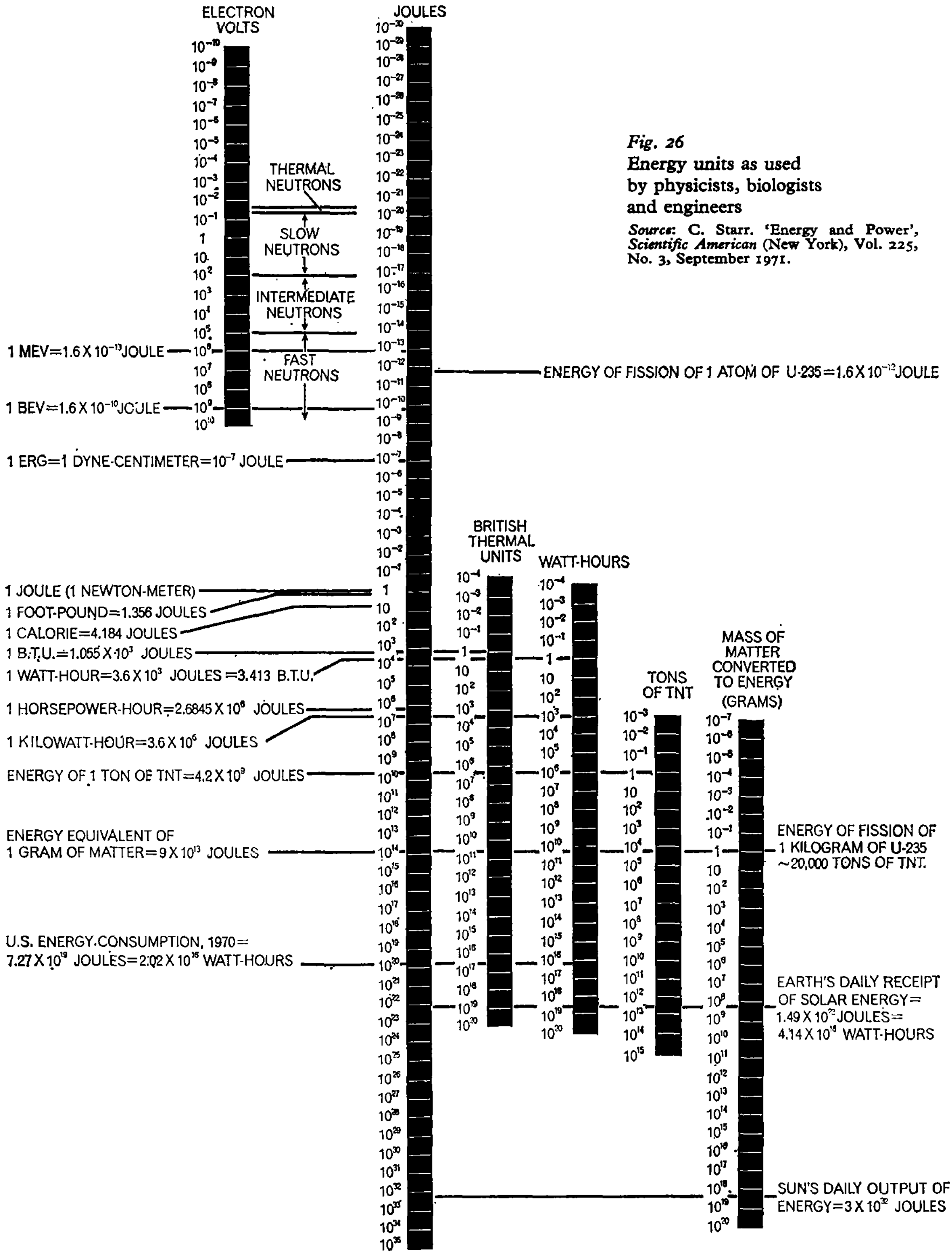


Fig. 26
Energy units as used by physicists, biologists and engineers

Source: **C. Starr. 'Energy and Power', *Scientific American* (New York), Vol. 225, No. 3, September 1971.**

than 30 per cent efficient in the conversion of chemical energy (fuel) into mechanical energy (motion); the remainder is lost largely as heat and through friction, which escapes from the system into the environment.

The second law of thermodynamics states that systems tend towards disorder. Physical and chemical processes proceed towards a maximally randomized state of energy. Such a randomized state of energy, which is available to do work, is defined as entropy. In other words, physical and chemical processes proceed towards maximum entropy in the system. Thus, increased order of any system is possible only through the expenditure of energy. Living systems (including man) overcome this tendency towards disorder only through the expenditure of more energy (generally through respiration) than is actually incorporated into the ordered structure of the organism.

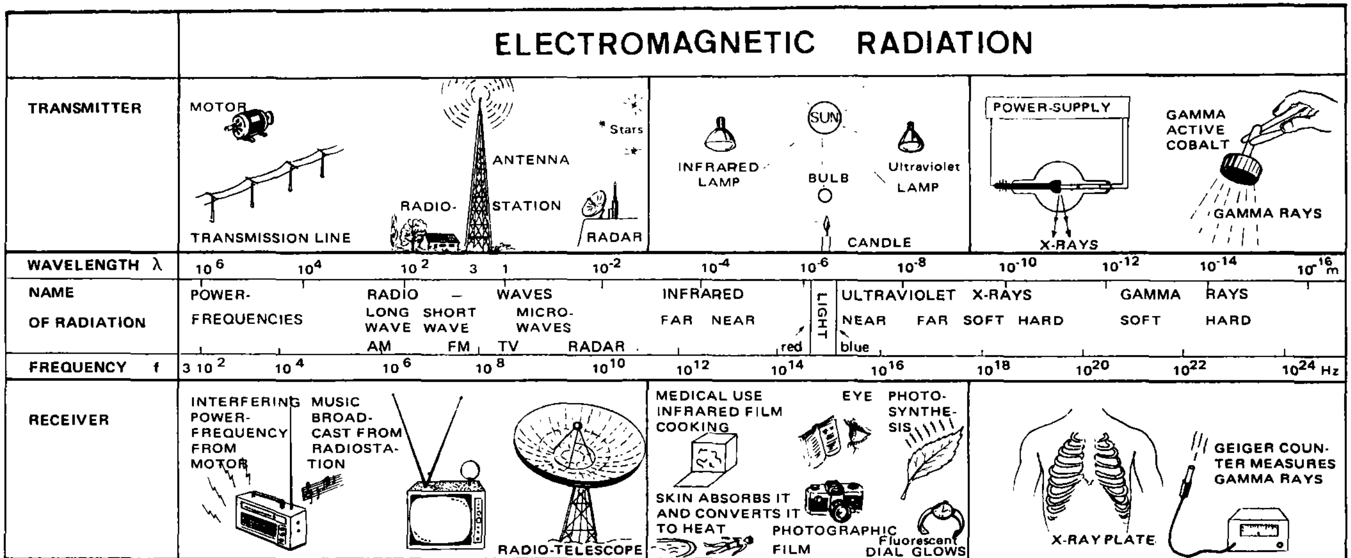
Living systems and energy

Every living system is but a particular manifestation of energy operating on a conversion and exchange basis within vital processes, namely in growth, differentiation, temperature maintenance (particularly in warm-blooded animals), mechanical work characteristic of muscle and cilia action, nerve and brain tissues eliciting electrical impulses, the remarkable phenomenon of 'active transport' in which substances are carried against osmotic gradients and in the processes generally grouped as metabolism. Living systems are, veritably, systems that are highly adapted for energy-transformation and energy-exchange. The major energy inputs to the earth are solar radiation, thermal energy from the planet's interior and tidal energy (which is an expression of gravitational energy). Of these, solar radiation is by far the greatest source of energy input and accounts for 99.98 per cent ($173,000 \times 10^{12}$ watts) of the total. About 47 per cent ($81,000 \times 10^{12}$ watts) of incident solar radiation is absorbed and converted into heat, 30 per cent ($52,000 \times 10^{12}$ watts) or more is reflected and scattered into space and 23 per cent ($40,000 \times 10^{12}$ watts) is used up in the hydrologic cycle. Less than 1 per cent is utilized by green plants and stored as chemical energy in food. A very small percentage of this energy has been stored in fossil fuels (coal and oil).

Solar radiation is a major part of the electromagnetic spectrum; it extends from the short-wave-length X-rays and gamma rays to the long-wave microwaves and radio waves (Fig. 27) but about 99 per cent of the total energy present occurs in the region of the wave-lengths 0.2 to 4.0 microns (i.e. the ultra-violet to the infra-red) and about one-half of this energy occurs in the visible spectrum of 0.39 to 0.76 microns.

The above facts suggest that solar radiation is the principal ultimate source of light, heat and chemical energy (in the form of foods and fossil fuels) on which human populations depend. Fossil fuels, such as coal, oil and natural gas constitute a major source of energy used by industry and commerce.

The basic principle of the quantitative convertibility of energy in living systems arises from the fact that the ultimate source of solar radiation is nuclear, i.e. arising from nuclear reactions taking place within the sun. The emissions which we refer to



DESCRIPTION

POWER-FREQUENCIES

In the W. I. most power transmission has a frequency of 50 Hz. This is a very small frequency, which implies a a very large wave-length of *about 6 000 km. Most of this* energy is confined to the power-lines and very little is radiated into space. Electric motors, spark-plugs in cars, etc. radiate waves which may interfere with radio and TV-reception, unless suppressors are used.

RADIO WAVES

They cover a large band of the electromagnetic spectrum. Wave-lengths range from about 30 km down to 30 cm in the very short wave radio bands.

MICROWAVES

They are used in long distance telephone calls and television transmission.

Radio astronomy uses special telescopes to receive microwave radiation from the stars.

Radar is one of the most important tools in modern navigation. It uses pulses of microwaves (wavelengths 1 — 10 cm) transmitted in a narrow beam by a parabolic mirror. If the beam hits an object — ship, areoplane, etc. — it is reflected back and detected at the radarstation.

INFRARED RADIATION

These radiations have wavelengths in the region 10^{-4} m. When absorbed by matter they are converted into heat and are therefore called "heat waves". All bodies emit infrared radiation all the time; the actual *wavelength depends on the temperature* of the body. The higher the temperature the shorter the wavelength.

LIGHT

Special cells at the back of the eye are sensitive to a very small range (4000 - 7000 angstrom) of electromagnetic radiation which we call light.

ULTRAVIOLET RADIATION

These radiations have shorter wavelengths than light. The most important properties are the ability to cause:

- chemical changes including action on photographic materials
- physical changes - fluorescence and phosphorescence
- biological changes - production of vitamin D in the skin; causes sunburn; used for sterilizing foods.

X - RAYS

X-Rays have very short wavelengths (10^{-8} down to 10^{-12} m). They are capable of penetrating matter and can therefore be used, with a photographic plate to insepct variations within a body e.g. bones within flesh.

X-rays also make certain materials fluoresce. X-rays help scientists find out about "how matter is built" - atomic structure and crystal structure.

GAMMA RAYS

Electromagnetic radiation with waves shorter than about 10^{-12} m is called gamma radiation. It is emitted when changes take place within atomic nuclei. Gamma rays have most of the properties of X-rays but are more penetrating.

Fig. 27
The spectrum of electromagnetic radiation

Source: H. Ibsedt. *Energy*. UWI, Unesco, Unicef, UNDP unit for Caribbean Teacher Training Colleges, 1973.

as electromagnetic radiation, come to the earth through space, their original being the nuclear reactions that occur in the solar furnace. On this planet, they are, in part, converted into chemical energy and, thence, by transduction in living systems they become the various vital forms of work, such as in biosyntheses (i.e. chemical work), transport and concentration (osmotic processes) and mechanical work. Uniquely, in any living system, free energy, which is but fractional in amount, is made available for work of any kind.

It must be remembered that living systems are isothermal, thus they operate at relatively low, constant temperatures and they do not convert heat energy directly into work, such as is done in an internal combustion engine or steam locomotive. Further, living systems, relying for their continuance upon energy-releasing reactions, are able to effect such releases by specific enzyme catalyses and controls, thus making possible osmotic, thermal, chemical, electrical and mechanical work, typically integrated in every life form.

Atmospheric influences of solar radiation

At the short-wave-length end of the spectrum, ultra-violet radiation has a high enough energy level to break chemical bonds of molecules of living things. Obviously, this is damaging to organisms and they must be protected from an excess of bond-breaking radiation. However, in the upper atmosphere, ultra-violet radiation converts oxygen (O_2) to oxygen atoms (O) and ozone (O_3) and ozone in turn strongly absorbs the ultra-violet radiation, and in this way the earth's surface and the human body are shielded considerably from harmful radiation.

Long-wave-length infra-red radiation is comparatively harmless to the human body, but it is of major significance. Radiation coming to the earth is of relatively short-wave-length and high energy value.

The greater proportion of short waves passing through the atmosphere and falling on the earth's surface are converted into radiations of longer wave-length, e.g. infra-red or heat waves, and re-radiated. Thus the planet both receives and conserves its heat by this process. The atmosphere acts generally as an energy-filter, allowing, for example, short-wave-lengths inwards and retarding the radiation of long-wave-lengths outwards.

The 0.032 per cent carbon dioxide in the atmosphere strongly absorbs radiation in certain infra-red bands of the electromagnetic spectrum. The clouds and water vapour in the atmosphere behave in a similar manner. Since radiation from earth into space occurs at the infra-red wave-lengths as dark-heat radiation, the effect of the above is that night temperatures are considerably warmer than they would be without such atmospheric absorption.

Other atmospheric influences by solar radiation include air circulations and their relationships to the distribution of water.

Environmental heat from solar radiation

Human heat balance

The bodies of warm-blooded animals are usually maintained at a constant temperature, with very little variation; such animals have body temperatures somewhat lower when sleeping than when awake and active. In humans, the variation is around 37° C. Muscles are by far the greatest producers of heat in the body, and, consequently the greatest consumers of energy. Skeletal muscles produce much more heat than do smooth muscles, because of their greater mass and because they can contract rapidly and vigorously. Shivering when in a colder, environmental temperature than usual, is an indication of the effective warming action of contracting muscles. The processes of heat production are intricate, involving the release of chemical energy from various assimilated food materials and several hormones, the chief of which is thyrosin; this hormone breaks off oxidative reactions from the synthesis of high-energy compounds.

The heat balance of the human body is influenced by three major physical processes: radiation, convection, evaporation. The body receives radiation from environmental objects with higher surface temperatures than the body, and it loses radiation to colder objects. Outdoors the body may receive radiation from the sun, the soil, a road surface or a building.

Similarly, if seated indoors near a cold window, the body is likely to lose heat by radiation to the window even though the room temperature may be warm enough to offset such a heat loss. Generally rough, dark surfaces radiate maximally while smooth light surfaces do so less effectively. The application of this principle to the colour of clothing is not entirely understood, but since light coloured clothing tends not to radiate as rapidly as dark clothing, it is often used in extremely hot environments because of less heat gain and also in extremely cold environments because of less heat loss. Lightweight clothing has been found to provide considerable protection from the sun with clothed subjects gaining 120 kilocalories an hour compared with 200 kilocalories an hour for unclothed subjects.

The conduction of heat to and from the body by air motion is influenced by convection. Heat is carried to or from the body depending on the temperature and movement of the air.

A reduction of blood flow near the surface of the body through vasoconstriction lowers the skin temperature and reduces heat loss to the environment. On the other hand, vasodilation increases peripheral blood flow, raises skin temperature to approximate core temperature, and increases heat loss to the environment. Similar processes also apply to radiant heat exchange.

As long as the atmosphere is not saturated by water vapour, there is some loss of heat from the body by evaporation. Over-all, the major mechanism here is the evaporation of sweat and through breathing. In both cases, the lower the humidity of the air at a given temperature, the greater the possible heat loss by evaporation is likely to be since the evaporation rate depends upon temperature. Considerable heat energy is required to convert water from the liquid to the vapour state. This is called latent heat

of vaporization, and amounts to 0.58 kilocalories for 1 ml of water. Excluding the process of sweating, about 900 ml of water are lost from the human body per day by urine and faeces in temperate zones, plus 400 ml from the lungs. This amounts to a heat loss of 754 kilocalories per day plus any thermal sweating that may be occurring.

Radiation, convection and evaporation are the major physical processes accounting for gain and loss of heat by the human body. Other processes also occur but they amount to a much smaller proportion of the total heat gain and loss and will not be discussed here. (See Table 7.)

Table 7
Heat balance

Factors increasing heat production (over basal metabolic rate)	Factors enhancing heat loss	Factors decreasing heat loss
Exercise or shivering Imperceptible tensing of muscles Chemical increase of metabolic rate Specific dynamic action of food Disease (fever)	Sweating Panting Cooler environment Increased skin circulation (vasodilation) Decreased clothing or shorter fur insulation Increased insensible water loss Increased radiating surface Increased air movement (convection)	Shift in blood distribution Decrease in tissue conductance Counter-current heat exchange

Source: G. Edgar, Jr. Folk. *Introduction to Environmental Physiology*. Philadelphia, Lea and Febinger, 1966, p. 83.

In general, then, the metabolic heat made available through the energy release of foods in the body provides for temperature equilibrium if metabolic heat input and the heat loss are balanced.

Because of this, a relationship is generally established between internal body temperature and that of the immediate environment.

Relations of heat to human activities

Without a doubt, man shows remarkable adaptation to a wide range of environments and in this section these adaptations are explored especially in relation to extremes of temperature.

For example, poorly clad people such as the Australian aborigine, the Alacaluf Indians (Tierra del Fuego) of southern Chile, the Bushmen of the Kalahari desert (Africa) and the sea-diving women of Korea, emphasize the point that man has been

able to make remarkable adaptation, both animate and technological, for coping with cold rather than with heat.

With appropriate use of shelter, fire and clothing, man has been able to accommodate himself to extreme cold conditions. In the simplest state, a shelter may be just a windbreak such as a Touareg tent, which is quite adequate for survival in the desert. On the other hand in some colder climates, more sophisticated shelters with central heating systems are provided. From the earliest civilizations, fire has been a major source of heat energy for comfort and food preparation, and wood has been the potential natural resource together with other plant material, e.g. peat.

Today coal, oil and gas have become the predominant fuel source for fire, and their use has extended beyond basic survival needs to such functions as very carefully controlled artificial climates, industrial power plants and transportation. Electricity and nuclear energy are becoming increasingly prominent additional means of meeting these needs. These benefits are not without cost, however, since the energy requirements increase as the energy services demanded increased. Implications of this fact concerning human ecosystem management are considered in a later chapter.

Clothing can be a major factor for comfort and survival, and the early invention, using animal skins worn outside-in was quite effective. The skin provided impermeability to the cold wind and the fur provided insulation. Heaviness, bulkiness and the difficulty of providing for evaporation of sweat during hard work are limitations to solving cold temperature problems through clothing.

It seems that much less can be done easily to solve problems of extreme heat. Lightweight clothing, as has already been pointed out, is quite effective in reducing direct radiation from the sun. This is beneficial under conditions of very high solar radiation since direct sunlight may add a heat load of 250 kc per hour. In many other cases, the benefit may be more than offset by the reduced evaporation and restricted convection caused by the clothing.

Since much more metabolic heat is produced in physical work than at rest, more heat dissipation will be required. An obvious response under conditions of environmental heat stress is to rest in the shade, but the demands of modern technological progress often limit feasibility of this approach. Furthermore, if the society is to be highly productive, these needs must be met when and where they occur, often placing much greater heat stress on the individual than he would experience in a more primitive condition. These problems are made more complex by clothing particularly under humid conditions, since the effectiveness of the main cooling mechanism of the body (sweating and evaporation) becomes limited. Heat disorders may lead to fainting, heat oedema and fatigue. Another effect of excessive heat is salt depletion from sweating, resulting in nausea, muscle cramps, circulatory failure and dehydration. Death from heat disorders is likely to occur at a body temperature of 40.5 to 42° C with the very young and relatively old individuals being the most susceptible to such heat stroke. During the summer of 1972, a severe heat wave in northern India resulted in 800 human deaths, illustrating the significance of heat stress on population; in 1973 in West Africa, severe heat and drought caused the death of thousands of humans, domestic animals and plants.

However, there are other effects as well which may be less obvious but are highly significant.

Efficiency in performing physical skills and mental tasks is often also related to temperature. Studies have shown lower production rates and lower quality of factory work when temperatures are too high or too low. Industrial accidents have been shown to increase at air temperatures above 24° C and below 12° C. Typists are more error-prone under uncomfortable environmental conditions and telegraphers have been shown to make more mistakes when temperatures are above 32° C. Similar differences in progress and standardized test scores have been found between children

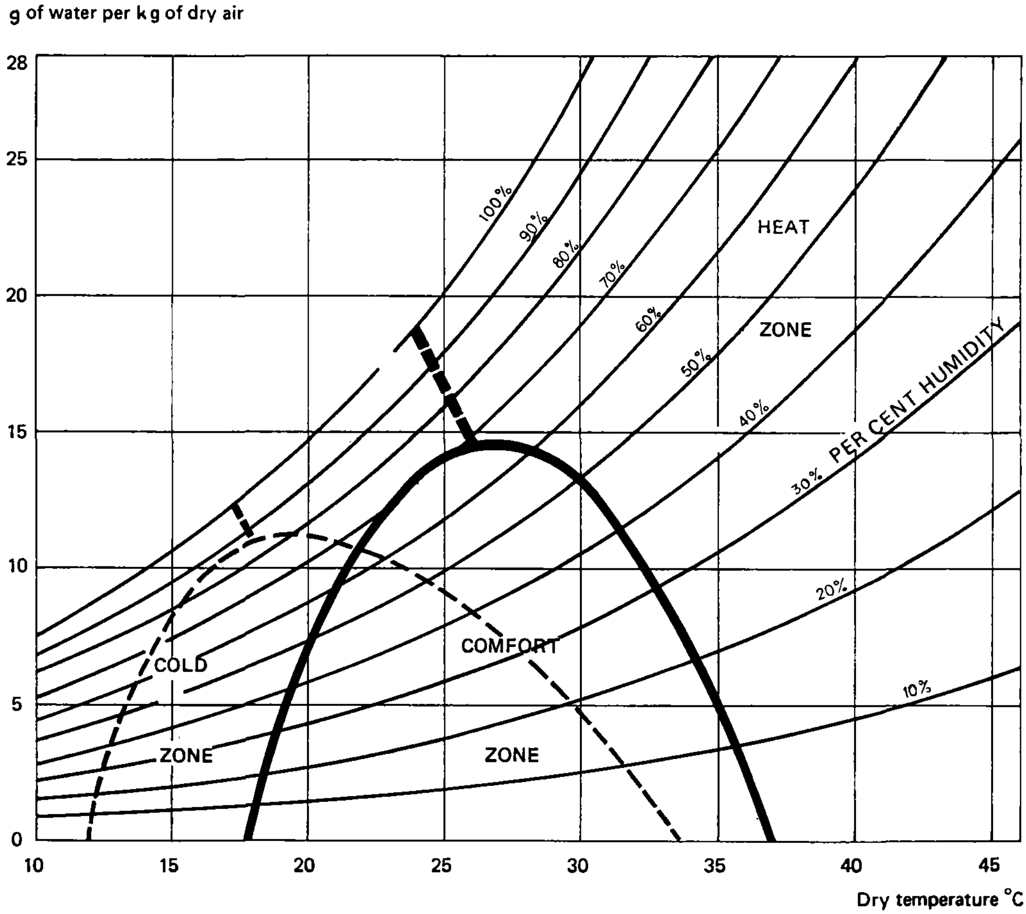


Fig. 28
Determination of the comfort zone for humans according to temperature and to air humidity

Source: R. Lemaire. 'Considérations physiologiques sur la climatisation en milieu désertique', *Journées d'information médico-sociale sahariennes*. Paris, PROHUZA, 1960, p. 101–12.

in climatically controlled classrooms (including air cooling) and those in classrooms heated when necessary but not cooled in hot weather. There may be many reasons for such differences in efficiency, but the fact that such differences exist is itself a matter of considerable consequence.

Various diseases have seasonal peaks, often in relation to disease carriers, such as insects, and in combination with temperature, as one of the significant climate factors. For example, typhoid fever and dysentery occur during the hotter seasons, with either increased incidence of greater frequency. Similarly, the frequency of respiratory infections tends to increase during the colder seasons, especially where climates are damp and cloudy.

In summary, the relation between heat and human activity is of significance in our consideration of environmental factors. Optimal temperatures for the performance of tasks can be identified, but the provision of these temperatures can seldom be accomplished by clothing alone. Artificial temperature control is the most suitable solution, but this demands extremely high quantities of energy whether for heating or cooling. The conflict is often increased by the fact that many human tasks in a technological society must be performed outdoors under conditions ordinarily considered too extreme for strenuous work. (See Fig. 28.)

Solar energy as food source

It has already been pointed out that a very small percentage of the solar radiation striking the earth is captured by green plants, converted to chemical energy and stored as food. It is this minute fraction of the solar radiation in which not only man but all living things depend for the continuation of their life processes as the ultimate source of chemical (food) energy. The process whereby green plants fix the energy of the sun is called photosynthesis. In this process, carbon dioxide (CO_2) and water (H_2O) are converted to glucose ($C_6H_{12}O_6$) and oxygen. The reactions take place in many steps, which will not be discussed here, but the summary of the over-all process can be stated as follows:

$$6CO_2 + 12H_2O \xrightarrow[\text{chlorophyll}+\text{enzymes}]{\text{radiant energy}} C_6H_{12}O_6 + 6O_2 + 6H_2O$$

There are many structural, physiological and environmental factors which determine the efficiency of photosynthesis in nature. Some of these are the temperature, availability of minerals and water, the intensity and duration of sunshine and the fine structure of the chloroplasts.

Solar input

The solar energy per unit area falling on the earth is called the solar flux and amounts to about two calories per minute per square centimetre. It is estimated that the earth receives approximately 1.3×10^{24} calories of solar energy per year (or 1.3×10^{21} kc). Of this amount, about 30 to 40 per cent is reflected back into space and does not greatly influence the surface of the earth. The other 60 or 70 per cent of the solar radiation amounts to roughly 8.1×10^{23} calories per year. When all the scattering, reflecting and absorbing processes are taken into account, only about 53 per cent of the solar energy entering the earth's atmosphere is available at the ground. Fig. 29 gives the distribution of solar radiation by wave-lengths reaching the earth.

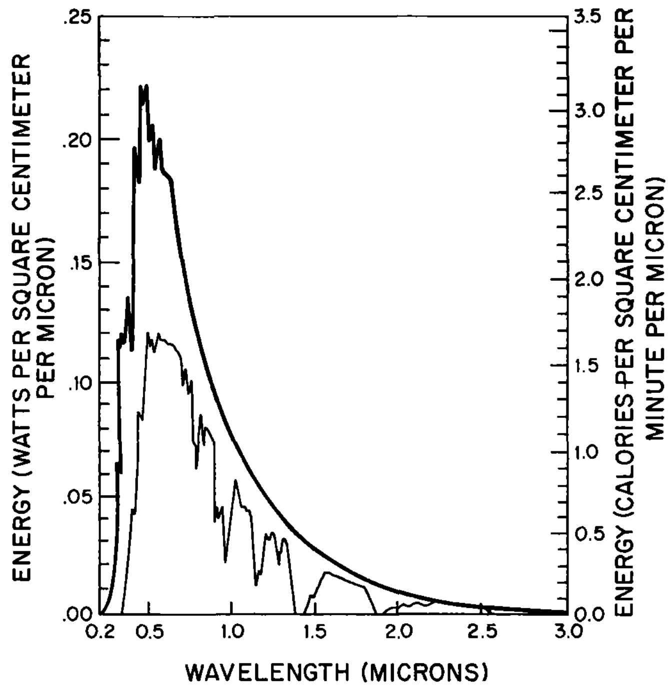


Fig. 29
Spectral distribution in upper atmosphere and surface of the earth

Source: David M. Gates. 'The Flow of Energy in the Biosphere', *Scientific American* (New York), September 1971, p. 7.

Note. Spectral distribution of solar radiation reaching the earth is given for the top of the atmosphere and the ground. Curve for ground takes into account the absorbing effects of water vapour, carbon dioxide, oxygen, nitrogen, ozone and particles of dust. Data are based on a solar constant of 1.95 calories per square centimetre per minute.

The solar radiation at the ground (direct sunlight plus skylight) varies from about 70 kilocalories per square centimetre per year in polar regions to about 200 to 220 kilocalories per square centimetre in desert areas. Tropical rain forests receive around 120 to 160 kilocalories while much of Europe receives around 80 to 120 kilocalories. Only about 25 per cent of the sunlight reaching the ground is of wave-lengths that stimulate

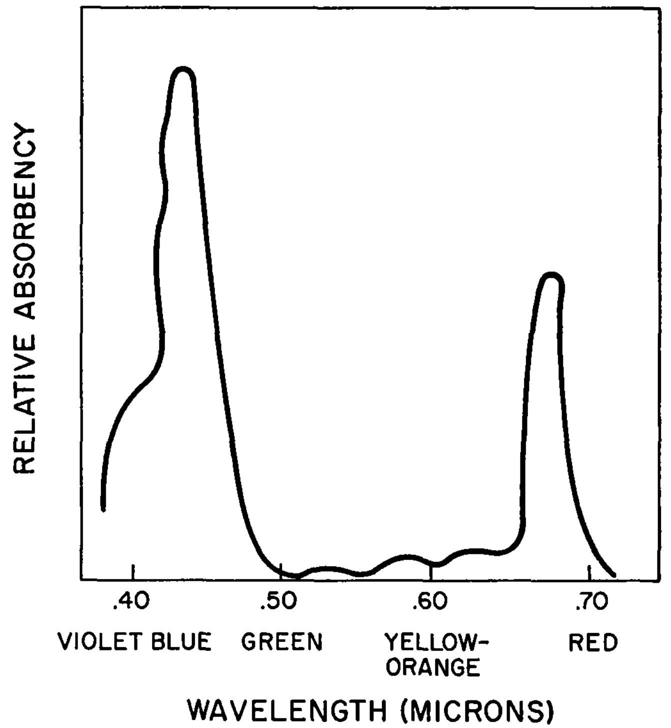


Fig. 30
Absorption spectrum of chlorophyll *a* in alcohol.
A note of caution: There is evidence that absorption of green light by the intact leaf—as it occurs in photosynthesis—is substantially higher than it is in an alcohol solution of chlorophyll

Source: William T. Keeton. *Biological Science*, second edition. W. W. Norton and Company, 1972, p. 104.

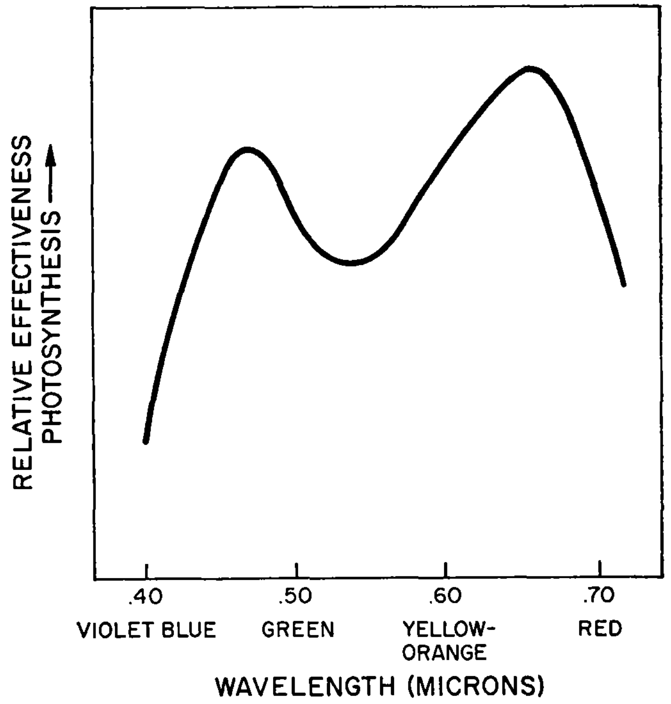


Fig. 31
Action spectrum of photosynthesis. Light of intermediate wave-lengths is more effective in driving photosynthesis than would be predicted on the basis of the absorption spectrum of chlorophyll *a*. Apparently, other pigments absorb these intermediate wave-lengths to some extent and pass the energy to chlorophyll *a* for photosynthesis

Source: William T. Keeton. *Biologica Science*, second edition. W. W. Norton and Co., 1972, p. 105.

photosynthesis, and only a fraction of this 25 per cent is actually used by green plants. Fig. 30 gives an absorption spectrum for chlorophyll. Fig. 31 gives an over-all action spectrum of photosynthesis.

Productivity

On the average, throughout the world about 320 g/dry weight (1.28×10^6 calories) of green plant matter is produced per square metre per year, or about 1.6×10^{16} g/dry weight (4.8×10^{20} calories) per year for the whole earth. This figure represents net primary productivity and does not include the food material which was metabolized by the plants in their own life processes. Essentially much of this net productivity is consumed annually by man, other animals and organisms concerned with decomposition. The energy not consumed shows up as increased biomass (tissues such as wood and muscle), humus and organic sediments.

Estimates of over-all efficiencies of land surfaces as a whole are round 0.1 to 0.3 per cent with estimates in water under natural conditions being fairly similar. It is possible to estimate the efficiency of net plant production calculated on input of solar energy during the growing season. Based on 500 calories per square centimetre per day of energy input, potential net plant production (gross production minus respiration) is about 71 grammes per square metre per day. This converts to 26.6 calories per square centimetre per day of solar radiation which results in biomass, or 5.3 per cent of the total incident solar radiation and 12 per cent of the energy received as visible light (222 calories per square centimetre per day). Field studies utilizing Sudan grass and also barley have confirmed efficiencies in this general range as possible.

The following examples further develop the comparison between the productivity of natural plant communities and agricultural crops developed for high productivity. Studies have shown that, on an annual basis, the net productivity in grammes per square metre per day may be around 9.0 for a spartina grass salt marsh at about 30° latitude N., 6.0 for a pine forest during years of rapid growth at 50 to 60° latitude N., 3.0 for a deciduous forest at 50 to 60° latitude N., 1.22 for a tall grass prairie at 40° latitude N., 0.19 for a short grass prairie at 40 to 50° latitude N. and 0.11 for a desert at about 40° latitude N. This compares with a potential net plant production of 71 grammes per square metre per day at 500 calories per square centimetre per day of energy input, and an estimated potential net photosynthesis of about 30 grammes per square metre per day at about 40° latitude N., during the four-month and eight-month growing seasons. At 40 to 50° latitude N., the annual value ranges from about 17.5 grammes per square metre per day to about 22 to 25. Fig. 32 shows net productivity for a number of natural and agricultural ecosystems.

The total net productivity (including stems, roots, leaves, fruits and grains) of the most productive agriculture is around 6,000 to 10,000 g/sq. m/yr. though the net production in most agriculture is around 1,000 to 3,000 g, in the same general range as many forests. This includes the above examples of corn and rice, though the yields drop far below these figures (to about 350 to 500 g), when expressed as grain production and across large areas.

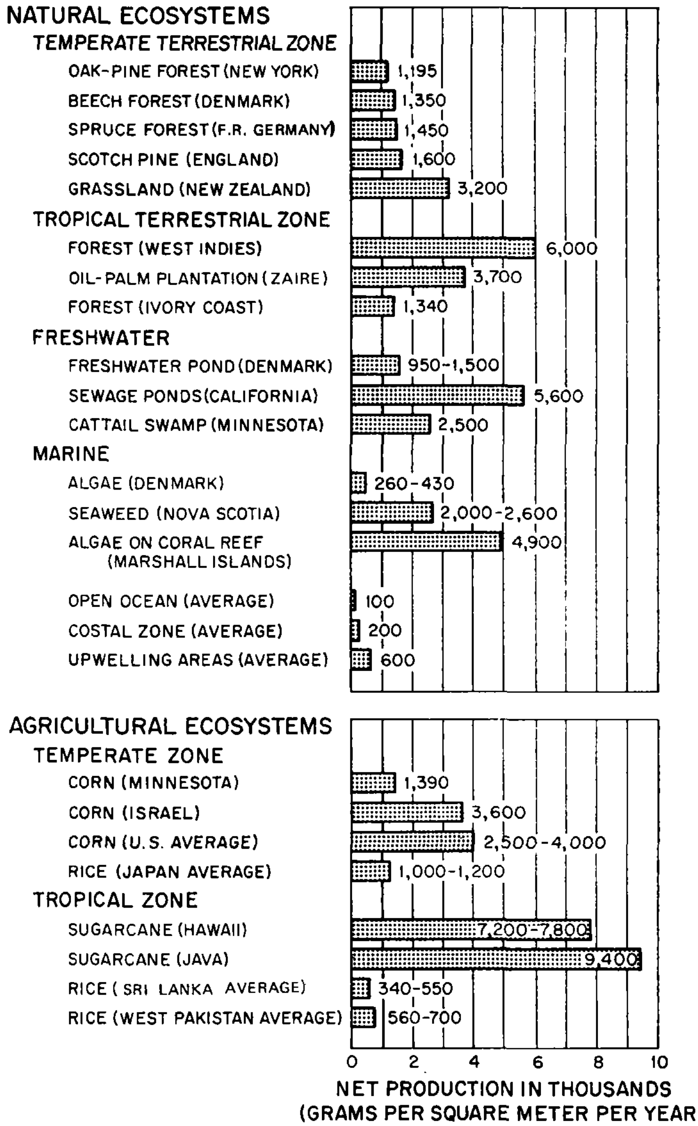


Fig. 32
Net productivity in various ecosystems

Source: George M. Woodwell. 'The energy cycle of the biosphere', *Scientific American* (New York), September 1970, p. 12.

Note. Net production levels of a number of natural and agricultural ecosystems are compared. The total net production of U.S. corn and rice is calculated from grain yields.

Some natural ecosystems are highly productive and others conspicuously unproductive. Algae on a coral reef at 10° N. show net annual productivity of 4,900 g and a forest at about 20° N. about 6,000 g. Tall grass prairies at 40° N., by contrast have a net productivity of about 450 g and short grass prairies at 40 to 50° N., less than 100 g. The open ocean also falls into this comparatively low productivity group. It appears that the ocean does not represent a vast, unexploited food reserve.

Net productivity for agricultural ecosystems (in g/sq. m/year)

Crop	Japan	China[1]	India
Corn (maize)	275	248[2]	96
Rice	192	276[2]	170
Wheat	207		121
Barley	253		98
Sugar-cane	5 840	6 410	4 910
Potato	2 270	840[2]	790
Sweet potato	1 990	840[2]	700
Dry peas	116	94	78
Dry beans	142	70	29
Ground-nuts	207	126	83
Rape seed	157	36	49
Sesame seed	50	41	23
Linseed	50	58[2]	26

1. FAO estimate data.
2. 1961–65 average.

Source: Summary of major crops, 1970. FAO Production Year Book, 1971.

Based on about 4,000 calories per g of dry organic matter produced, efficiencies of production of the natural plant communities can be estimated. For the salt marsh, the figure is 0.8 per cent, the forest is about 0.5 per cent efficient, the tall grass prairie 0.1 per cent and the desert about 0.05 per cent. A cornfield at 40° latitude N. can be expected to be about 1 to 2 per cent efficient.

It was suggested earlier that various factors (such as light intensity, light duration, temperature, availability of water, soil conditions, availability and concentration of various minerals, and concentration of carbon dioxide) influence net productivity by affecting photosynthesis, respiration and the over-all condition and health of the plant. While optimum conditions are not identical for all species of plants in all regions of the earth, there is a general correspondence of actual net productivity to such factors as mineral resources, hours of sunlight per day during the growing season, optimum temperature, and availability of water. Highly productive ecosystems (3,000 g/sq. m/per year net production, and above) are areas like river estuaries and their dependent salt marshes, coral reefs, flood plain forests, paddy fields and sugar-cane plantations. Such ecosystems generally exist in warm-temperate regions and have an abundant water supply and considerable raw material resources. Ecosystems of average net productivity (1,000 to 2,000 g/m^2/yr.) are temperate forests, many agricultural crops, shallow lakes and continental shelves. Relatively unproductive ecosystems are deserts, open seas and lakes, tundra, grasslands, shrublands, open woodlands and some cereal crops. Restricted productivity of these is often due to limited water, lower raw materials levels and restricting temperatures.

From the standpoint of energy resources and efficiencies of ecosystems, the first major loss of food resources is in respiration by the green plant itself and this is the

factor which accounts for the differences between gross productivity and net productivity so far discussed.

The actual percentage of the gross productivity consumed in respiration by the producers of the ecosystems varies widely not only with environmental factors but also with the structure of the ecosystem. For example, young ecosystems and agricultural ecosystems need not support a large mass of living tissue, as must a mature forest. The forest tends to consume a greater proportion of the gross productivity than the agricultural crop (though net productivity of the forest may be quite high because of a very high initial gross productivity). Some studies have reported as little as 15 to 20 per cent of the gross production consumed in respiration by the producer, while in other studies the figures run as high as 70 to 75 per cent. These highest figures are more typical in tropical forests with a lower range of 50 to 60 per cent in temperate forests. Over all, green plant respiration may be considered to account for about 30 to 40 per cent of gross production, so that net production is about 60 to 70 per cent of gross production.

There are other energy leaks prior to gross production, resulting in an efficiency of energy capture of roughly one per cent of the total incident radiation. Using the estimates suggested above, we have net productivity equal to about 0.6 to 0.7 per cent of total incident radiation, and often less. This is the approximate efficiency with which total incident radiation is stored in green plants.

Trophic levels and food webs

On a world-wide basis, almost all of the energy fixed and stored by the producers is consumed by other living organisms. Much of that which is not consumed by herbivores and omnivores is consumed by decomposers. The several nutritional levels from producers to herbivores to primary and secondary carnivores and decomposers are jointly referred to as trophic levels. It is also possible to describe the specific

Table 8
Marine and terrestrial food chains[1]

Trophic level	Marine food chain	Terrestrial food chain
Producers	Phytoplankton	Grasses
Primary consumers (herbivores)	Herbivorous zooplankton	Locust, antelopes
Secondary consumers (1) (general carnivores)	Sardine	Birds Smaller mammals (genet cat)
Secondary consumers (2) (intermediate carnivores)	Herring	
Tertiary consumers (3) (top carnivores)	Tuna	Larger mammals (e.g. hyena, lion)

1. As is frequently the case, the marine food chain contains more steps than the terrestrial example. One explanation advanced for this common difference is that marine systems are older and have had a longer period in which to evolve.

Source: Arthur S. Boughey (ed.). *Fundamental Ecology*. San Francisco, Intext Educational Publishers, 1971.

nutritional relationships of a community in terms of the species involved. Such a series of specific relationships is a food chain; Table 8 shows an example of a marine food chain and terrestrial food chain.

In actual communities of living things, there are ordinarily numerous species identified with each trophic level. Not all organisms at one trophic level feed on the same species of the next lower trophic level, of course. Rather, numerous simple food chains (as illustrated in Table 8) operate simultaneously in a community and become intricately interconnected to form a 'food web'. Fig. 33 provides an example of a simple food web.

A knowledge of the precise inter-relationships of the food web of a given ecosystem is needed if man is to be aware of the full implications of any planned ecosystem modifications. Of more immediate concern here is the energy utilization efficiency at the various trophic levels.

The net productivity represents the food produced beyond the consumption of food by the producer. If the herbivore feeds on this food (of net productivity), it will convert some of the food material into living tissue or biomass of its body, but will consume much of the food for energy in the process of respiration. A similar situation exists for the carnivore, the secondary carnivore and so on. If we now view these trophic levels (producer, herbivore, carnivore, etc.) as potential food (or energy)

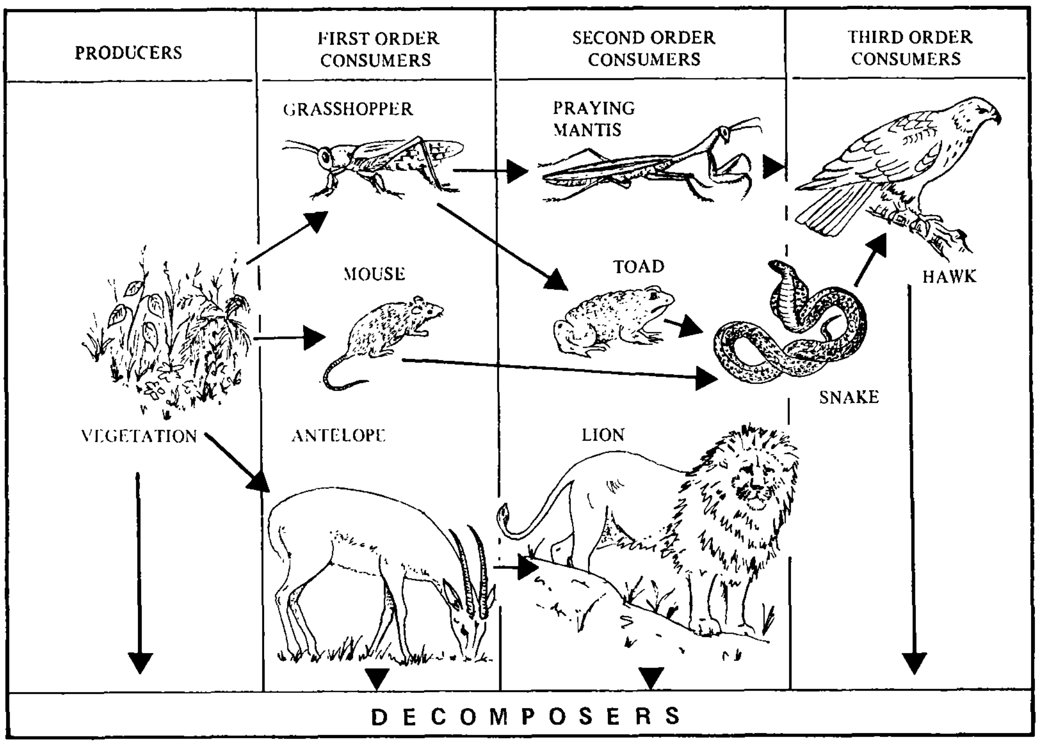


Fig. 33
A simple food web

Source: Adapted from ***Investigating Living Things. Animals as Food Consumers.*** Unesco Pilot Project for Biology Teaching in Africa, 1968.

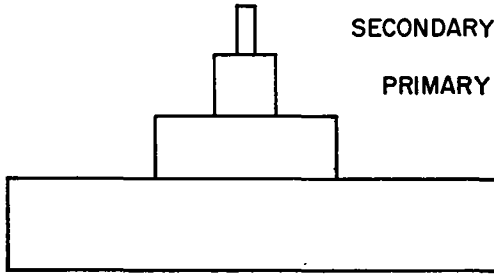


Fig. 34
Energy pyramid

sources for man, it is apparent that the efficiency of total energy available as a food source at each successive trophic level decreases. The over-all quantitative energy relationships among the trophic levels are often depicted as energy pyramids. Fig. 34 shows such an energy pyramid. The relative area shown for each trophic level in an energy pyramid represents the relative quantity of energy incorporated as food materials in that level. As would be expected, the actual values at each trophic level vary widely from ecosystem to ecosystem and from food chain to food chain within an ecosystem. Numerous studies along these lines have been made, however, and some estimates of efficiencies generated.

Net secondary productivity

Productivity in an aquatic ecosystem as recorded by the harvest method is in this particular instance net secondary productivity. The harvest method is just as applicable to the measurement of net secondary productivity as it is to the determination of net primary productivity. These figures illustrate the effect on productivity of adding nutrients and additional food into an aquatic ecosystem. They also show the reduction in total net secondary productivity in an aquatic ecosystem when a carnivorous fish trophic level is present.

Trophodynamic status of community	Net secondary productivity g/m² per year
Unfertilized waters	
North Sea, herbivore-carnivore	1.68
Great Lakes, herbivore-carnivore	0.9 – 0.8
Fish ponds, U.S.A. (with carnivores)	0.21– 18.1
Fish ponds, Federal Republic of Germany (without carnivores)	11.2 – 30.0
Fertilized waters	
Fish ponds, U.S.A. (with carnivores)	22.4 – 56.0
Fish ponds, Federal Republic of Germany (without carnivores)	99.7 –157.0
Fish ponds, Philippines (without carnivores)	50.4 –100.0
Fertilized waters plus added food	
Fish ponds, Hong Kong (without carnivores)	224.0 –448.0

Numerous energy factors and relationships operate in all ecosystems. Often there is importation of organic matter from other ecosystems and there may be indefinite storage of organic matter from any of the trophic levels and some is exported. Much of the energy efficiency loss from a lower trophic level to a higher one is due in part to energy consumed in respiration. Some of the energy remains unavailable to the organisms of the next level trophic because it is stored and can be exported to another ecosystem, or used by bacteria and fungi (decomposers) in the process of decay.

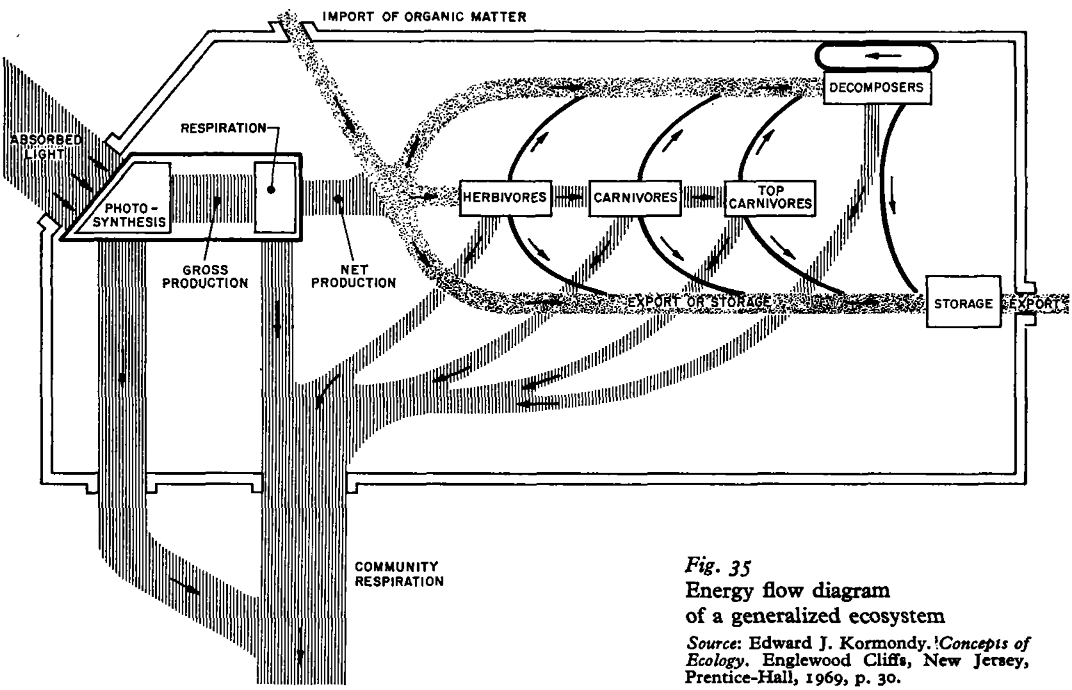


Fig. 35
Energy flow diagram
of a generalized ecosystem
Source: Edward J. Kormondy. *Concepts of Ecology*. Englewood Cliffs, New Jersey, Prentice-Hall, 1969, p. 30.

Fig. 35 shows these major energy relationships in an energy flow diagram of a generalized ecosystem.

Since man is also dependent upon the food energy incorporated from the incident solar radiation, the comparative efficiencies of the various food sources are of human interest. This is particularly true in relation to current human population growth and the existing inadequate diets of many such populations.

Human foods needs, sources, and efficiency

Since man is a consumer, the efficiency of his energy source will not be higher than for a herbivore, except as man is able to increase productivity through agricultural methods. Even here, an accurate accounting of energy relationships requires recording of the energy output through the agricultural methods. That is, the increase in net productivity is being generated through the expenditure of some energy, whether it be

the fuels and industry for the machinery, or primarily the increased physical activity and energy requirements of man (at the herbivore or carnivore trophic level).

It is estimated that the average daily *per capita* consumption of energy from food is about 2,200 kilocalories averaged over the entire world. Man's daily energy requirement is about 1,320 kilocalories per day if he is entirely sedentary, 2,400 kilocalories if he is moderately active and considerably more than this with higher activity and in colder climates. If we assume that the average basic energy needs of man are about 2,500 kilocalories per day, and if his entire energy requirement were to be met by eating herbivores, the requirement would be equivalent to about 25,000 kilocalories of herbivore, which would in turn require about 250,000 kilocalories per day of the producer. If the producer is about one per cent efficient this would in turn require about 25 million kilocalories per day of radiant energy. Based on solar radiation of 500 calories per square centimetre per day, it would require between 0.4 and 0.6 hectares to feed one person for one day. The 1.4×10^9 hectares under cultivation and 2×10^9 hectares of grazing land on earth provide the 3.5×10^9 people on earth with about 0.4 hectares *per capita* under cultivation and 0.6 hectares of grazing land. The total potential arable land is about 3.2×10^9 hectares and the total grazing land is about 5.2×10^9 hectares. While it is possible to subsist on around 2,000 kilocalories per day, it is not possible to be very energetic or function effectively in industrialized society at that level of energy intake. The implication is that under the above assumptions and estimates of efficiencies it would not be possible to feed more than about 2.5 times the present world population. In actual fact, the diet of the world human population is made up of food taken in at various trophic levels with a wide range of cultural and individual differences due to ecological, social, religious and personal preferences. But the advantage of using the producer level is obvious.

It is to be noted that this final comment is distinctive to man, as a highly adapted, socialized animal.

Land use

The preceding account given on human food needs clearly implies the importance of how land should be used, in order to obtain the maximal efficiency from the soil, the vegetation (both natural and man-made) and from natural faunas and domesticated livestock that feed upon the particular land areas. In proposing ways to increase food production, the increased use of cultivated or cultivable land is often suggested. Such ideas and proposals are discussed in Chapter 10 on 'food and nutrition'.

Industrial energy

Introduction

Today, we live in an age where the applications of science and technology, through the design and use of machinery in the factory and the home, are intrinsically bound up with the successful survival of the human population, with economic growth, physical

comfort and improved quality of life. These developments are largely determined by the sources and availability of energy resources and their profitable exploitation. Any such supplies of energy that run our machines and services, our technological innovations that go to support and develop our society, may be called 'industrial energy'. This conditioned use of energy is, principally, the result of exploitation of fossil fuels like coal, oil and gas. Other sources of industrial energy include power from water (either mechanical, electrical or chemical), from wind, from nuclear fission and from the ocean tides.

Historical review

Utilization of energy in modern society depends on two factors: available resources and technological skills (need for the exploitation and use of such resources). Human progress from 'Primitive Man' to modern 'Technological Man' is mainly a story of industrial inventiveness linked with ever increasing energy needs (Fig. 36). In this process, man has succeeded by various means (hunting, farming or by burning fuel) to introduce himself into the natural energy-cycle, converting it from less useful forms to more economically desirable ones.

Man's appearance as a hunter and food gatherer in the earth's system did not disturb the natural balance of input (photosynthesis) and output (respiration) of solar energy. At this stage, muscular energy was the only source for man's meagre needs. It was regularly replenished through his consumption of wild animals and plants as food. The idea and development of agriculture transformed most hunting populations into farming communities. Raising crops and domesticating animals afforded man some important means of exploiting the energy. The agricultural urban revolution, while still maintaining the basic pattern of conversion of food into muscular energy in the biosphere, effected two important changes: reduction in the high diversity and efficiency of the natural communities in order to store solar energy in species either edible or otherwise usable (i.e. draft animals) by man, and reduction in the flux shunting through life processes. At the same time, his investments in agriculture brought man higher dividends and enabled him to progress beyond bare existence towards modern complex system.

With the use of fire, primitive man took the first big step in the exploitation of an industrial energy resource. As he advanced further, he utilized water power by inventing the horizontal water wheel about the first century B.C. followed by the vertical water wheel by the fourth century, with a capacity to develop two kilowatts of power. The use of the windmill for grinding cereals, for hoisting materials from mines and for pumping water developed in the 12th century in Europe. At about the same time, man, in England, discovered some 'black rocks' that would burn. This discovery of coal as an energy source henceforth was followed by the mining of this material and its use for various purposes, including domestic heating and smelting of metals. With the advent of steam, two main resources of energy (namely, water and fossil fuels) were combined in the service of man and thus the second phase of the industrial

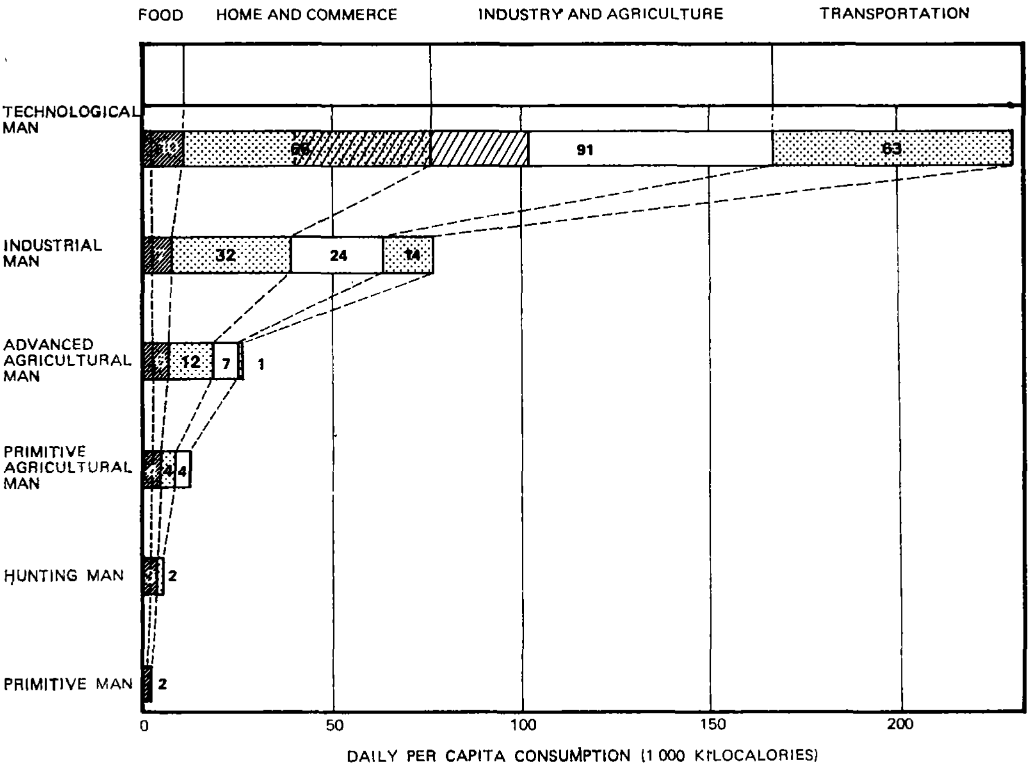


Fig. 36
Daily consumption of energy *per capita* for six stages of human development

Source: E. Cook. 'The Flow of Energy in an Industrial Society'. *Scientific American* (New York), September 1971.

Note. Daily consumption of energy *per capita* was calculated by the author for six stages in human development (and with an accuracy that decreases with antiquity). Primitive man (East Africa about 1 million years ago) without the use of fire had only the energy of the food he ate. Hunting man (Europe about 100,000 years ago) had more food and also burned wood for heat and cooking. Primitive agricultural man (Fertile Crescent in 5000 B.C.) was growing crops and had gained animal energy. Advanced agricultural man (north-western Europe in A.D. 1400) had some coal for heating, some water power and wind power and animal transport. Industrial man (in England in 1875) had the steam engine. In 1970 technological man (in the U.S.A.) consumed 230,000 kilocalories per day, much of it in form of electricity (*hatched area*). Food is divided into plant foods (*far left*) and animal foods (or foods fed to animals).

revolution was possible. Soon, there was development of the steam engine, locomotives, steamships and steam-electric power.

The fast expanding industry speculated on the discovery of natural gas by geologists as additional sources of fossil energy. With this came the inventions of the internal combustion engine, the automobile, the aeroplane and diesel-electric power. As a result, mobility was accelerated and agriculture was mechanized. The generation of electric power accelerated the process of modernization by the invention of the telephone, radio, television and transistor, etc. During the Second World War, the most significant source, nuclear energy, was brought under control; and, currently, this is fast becoming the world's greatest source of power in an expanded industrial civilization. (See Fig. 37.)

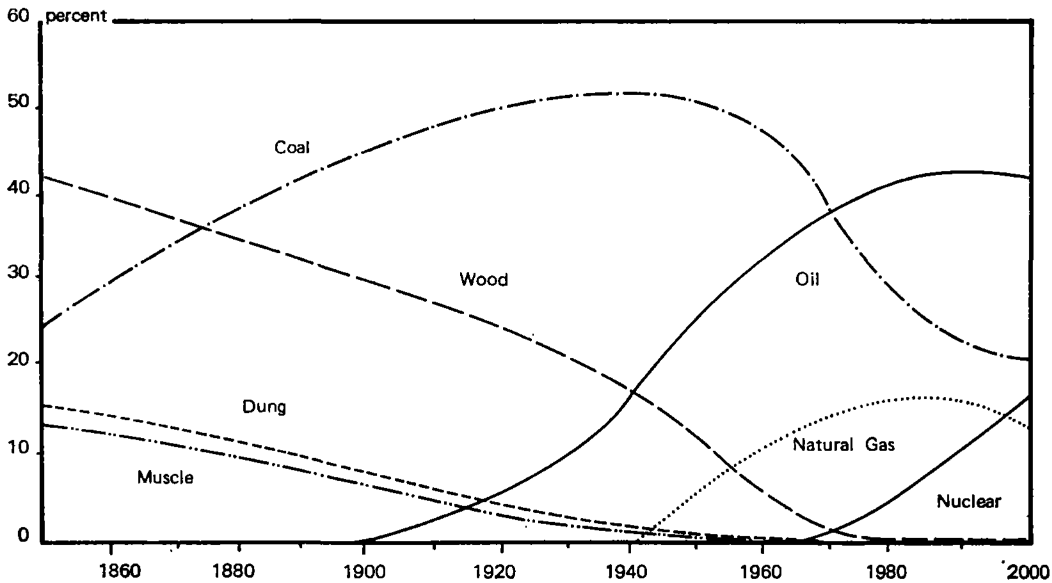


Fig. 37
World consumption
of energy sources
Source: J. Mchale. *World Facts and Trends*. London, Collier-Macmillan, 1973.

Sources of energy

Sources of energy are either sustained or exhaustible. The former does not undergo any appreciable loss in energy supply during use. Solar energy is one such source. In terms of power (solar radiation intercepted by one hemisphere) it is estimated to be 100,000 times as great as the total existing electric power. But the cost needed to make it practicably feasible is not economical. However, some direct use of this energy is made in solar batteries to generate electricity for space ships.

The mechanical energy of falling or moving water provides a second sustained source and has been developed to only about one-fifth of the total maximum potential available in the world. It holds great potential in South America and Africa where coal is in short supply.

The internal heat of the earth (giving geothermal power) is shown by the heat stored in hot springs and volcanoes and is another virtually limitless source of energy. It contributes towards the production of electric power at present in some countries (Table 9, page 105). Further developments are in progress for greater utilization of this source. However, except for hydro-electric power, all sustained sources of energy hold limited scope because of the very high cost of harnessing them.

The exhaustible resources include fossil fuels which have been and still are principal sources of energy. Over the years there have been great changes in the relative

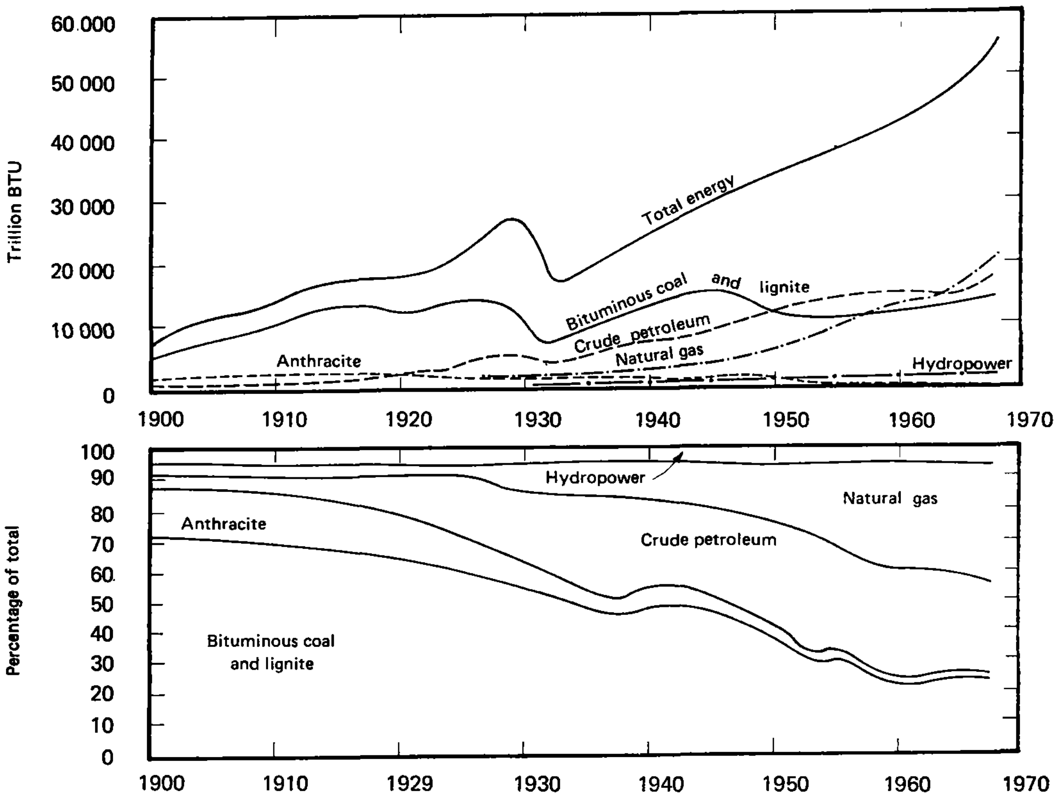


Fig. 38
Changes in production of fossil-fuel energy resources and hydro-electric power, 1900 to 1968.
Quantity of nuclear power is too small to show on this scale

Source: Data of U.S. Bureau of Mines, *Minerals Year Book*. From: A. N. Strahler. *Planet Earth: Its Physical Systems Through Geologic Times*. New York, Harper and Row, 1972.

contributions of each towards total energy supplies. For example, both petroleum and natural gas have increased in proportion since 1900. On the other hand coal and lignite have decreased to less than half of their starting percentage during the same period (Fig. 38). The decline in the coal component towards the total energy supply was not due to the fall in the amount of its consumption or production. Between 1929 and 1970 it has been estimated that world coal production increased by 1.7 times as against 10 times in the case of petroleum and 20 times in the case of natural gas. Both oil and gas are today's most favoured fuels. A brief description of world distribution and production of these fuels is given below:

Table 9
Geothermal power development

Country	Capacity (thousands of kilowatts)
Iceland	17
Italy	390
Japan	33
Mexico	4
New Zealand	170
U.S.A.	83
U.S.S.R.	3
TOTAL	700

Source: J. B. Koenig. *Geotimes*, Vol. 16 (3), p. 12.

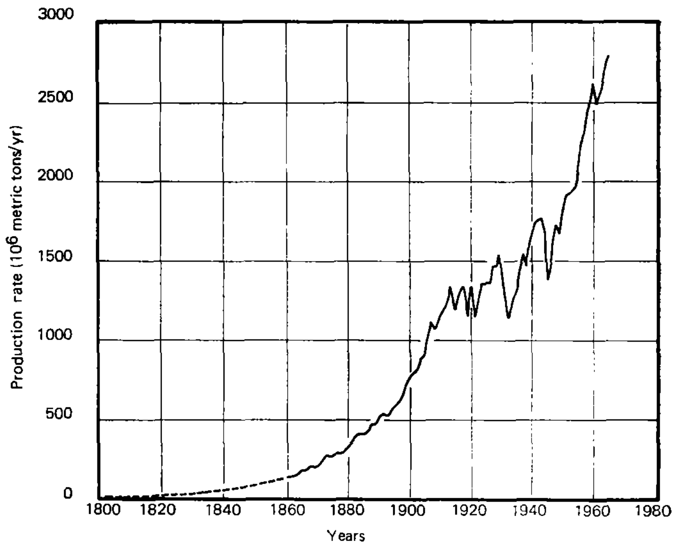


Fig. 39
World production of coal and lignite
Source: Resources and Man: a study and recommendations by the Committee on Resources and Man, National Academy of Sciences—National Research Council. San Francisco, W. H. Freeman and Company, 1969, p. 161.

Production of coal

Despite the development of other fuels and that of power industries in the present century, coal is still vital and perhaps the main source of energy in many countries. Leading coal-producing countries in the world include China, the U.S.S.R., the U.S.A., the Federal Republic of Germany, the German Democratic Republic, the United Kingdom, India, Japan, Poland, Czechoslovakia and France. Fig. 39 shows the rise in the world's annual production of coal and lignite since 1860, while Table 10 lists production, by world regions for the year 1970.

Production of oil

The major oil-producing countries include the U.S.A., Venezuela, the U.S.S.R., Kuwait, Saudi Arabia, Iran, Iraq, Nigeria, the Libyan Arab Republic and Indonesia.

Table 10. Production of fossil fuel and energy in 1970[1]

Region	Coal (million metric tons)	Lignite (million metric tons)	Natural gas	Crude petroleum (million metric tons)	Energy (million metric tons) of coal equivalent	Electrical energy (millions kWh)
World	2 123.0	792.0	1 071.0	2 278.0	7 000.0	4 900.0
Africa	59.6	—	3.1	294.3	450.0	87.8
America, North	554.8	8.9	702.8	566.9	2 331.1	1 897.3
America, South	7.4	0.1	21.0	239.3	358.8	107.0
Asia	520.9	16.2	31.1	779.9	1 604.3	597.1
Europe	501.8	596.3	113.9	37.0	974.6	1 398.7
Oceania	46.0	26.1	1.6	8.5	71.3	72.1

1. World aggregates include the U.S.S.R. which is not shown separately nor included in the continental totals for Asia and Europe. Owing to the rounding of figures, the world aggregates do not always correspond exactly to the sum of the continental breakdown.

Source: U.N. Statistical Year Book, 1972.

Fig. 40 shows annual world production of crude oil since 1880. Fig. 41 gives production and consumption by region for 1970.

Next to the oil-rich Middle East countries, Nigeria and the Libyan Arab Republic are the largest producers of oil in Africa.

Production of natural gas

Natural gas supplies 25 per cent of United States fuel energy. The natural gas-producing countries are mainly Canada, the U.S.A., Mexico and Algeria. However, in the Middle East, Venezuela and Indonesia, it occurs in association with oil.

Electrical energy

Modern industry relies primarily on electrical energy, which is derived from primary sources such as oil or water. Electricity can be transmitted to the consumer and can

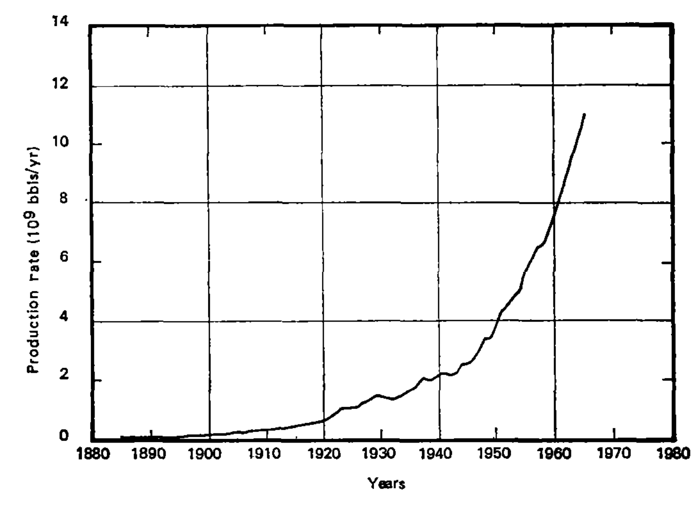


Fig. 40
World production of crude oil

Source: Resources and Man: a study and recommendations by the Committee on Resources and Man, National Academy of Sciences—National Research Council. San Francisco, W. H. Freeman and Co., 1969, p. 162.

easily be reconverted by him for heating, lighting or the production of mechanical energy to drive machines. Its average annual increase in production over the past decades has been very high even in countries with a modest level of economic development, such as in Zambia, Nigeria, Morocco, Kenya and India.

Per capita *energy consumption*

Industrialization processes today are characterized by two trends: enormous consumption of energy which is steadily increasing and total dependence on fossil fuels. As a result, global *per capita* energy consumption has more than doubled during the

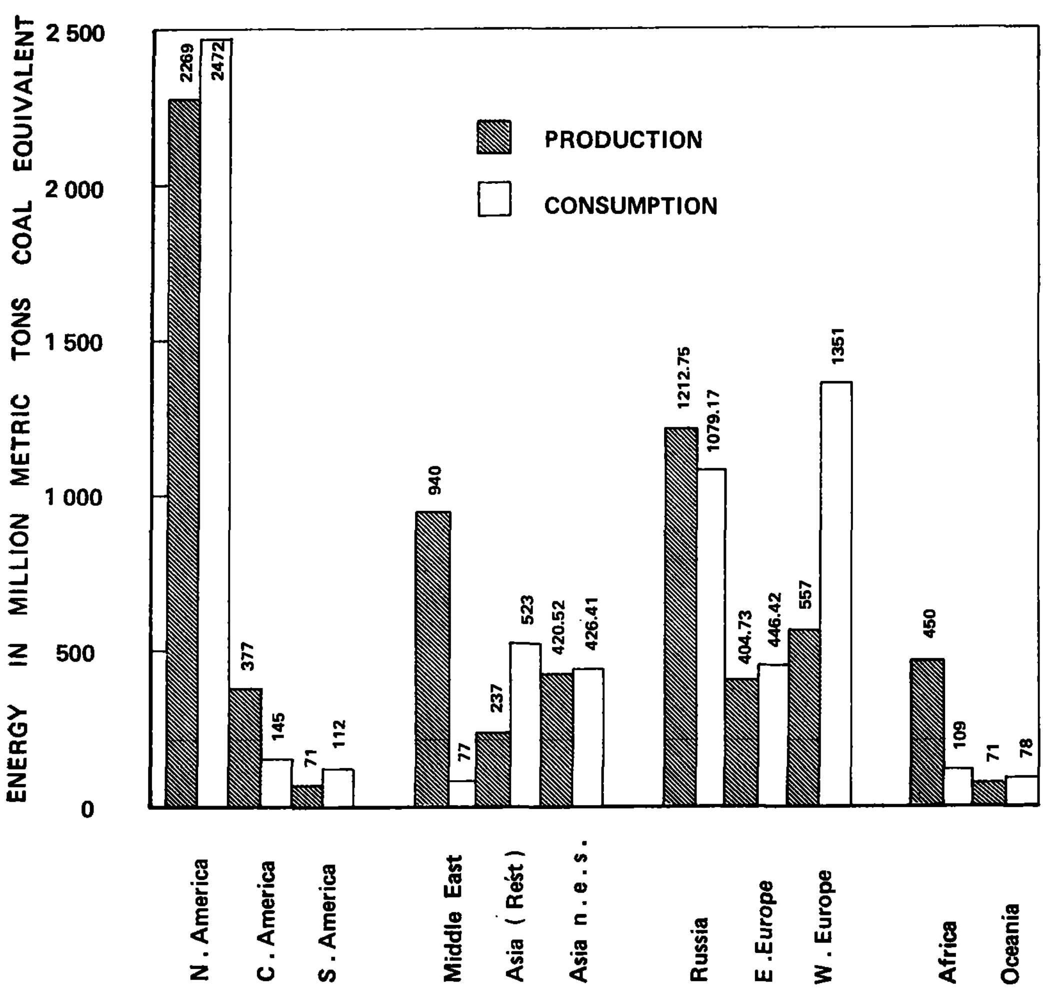


Fig. 41
Production and consumption of energy by region (1970)
Source: UN Statistical Year Book, 1972.

Table 11. Global consumption of primary sources of energy

Year	Total energy		Primary sources of energy (in percentage)			
	Aggregate In million metric tons coal equivalent	*Per capita* In kilogrammes coal equivalent	Solid fuels	Liquid fuels	Natural gas	Hydro and nuclear electricity
1929	1 713	867	76.0	14.1	4.3	5.6
1937	1 826	900	70.0	17.0	6.0	7.0
1950	2 519	1 054	58.5	23.8	10.2	7.5
1960	4 233	1 403	52.1	31.2	14.6	2.0
1970	6 817	1 889	35.0	41.8	20.8	2.3

Source: U.N. Statistical Year Books, 1950, 1960 and 1972.

last four decades (1929 to 1970) and total energy consumption has increased about fourfold (Table 11).

Energy consumption *per capita* is closely related to the general level of development; the highly developed industrialized economies have by far the largest *per capita* consumption. For example, each U.S. citizen is responsible for burning the equivalent of more than 11,000 kg of coal every year; and likewise, each Czech citizen, over 6,000 kg, each British citizen over 5,000 kg. As shown in Fig. 42, these rates contrast greatly with those in the developing countries, say India where *per capita* energy is 191 kg equivalent of coal per annum. Moreover, its rate of increase in India, within the last two decades, is also insignificant.

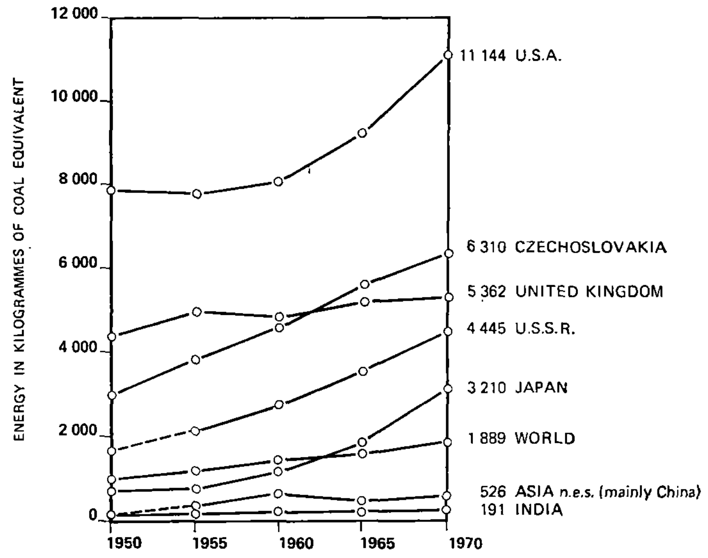


Fig. 42
Per capita energy consumption since 1950
Source: U.N. Statistical Year Books, 1950, 1960, 1972.

An Ethiopian nomad, herding his sheep over sparse eastern African range land, consumes only about 34 kg equivalent of coal in a year. Fig. 36 summarizes daily consumption of energy *per capita* for six stages in human development. Primitive man (1 million years ago) used energy of food alone, hunting man (1 million years ago) had more food and also used fire for heat and cooking; primitive agricultural man (5000 B.C.) grew crops and domesticated animals for energy; advanced agricultural man (A.D. 1400) acquired fossil fuel, wind- and water-power for his needs; industrial man (from 1870) added steam power to his energy resources and the technological man (from 1970) depends principally on electric power and consumes about 230,000 kilocalories per day for his various machines and needs. Thus the affluent man in technological society has become dependent upon ever increasing amounts of energy. For example, consumption is doubling every ten to fifteen years in the United States and in the European community it has trebled over the last twenty-one years. There is close relation between a nation's consumption of energy *per capita* and its gross national product (Fig. 43). Comparative data on world energy production and consumption are represented in Fig. 44.

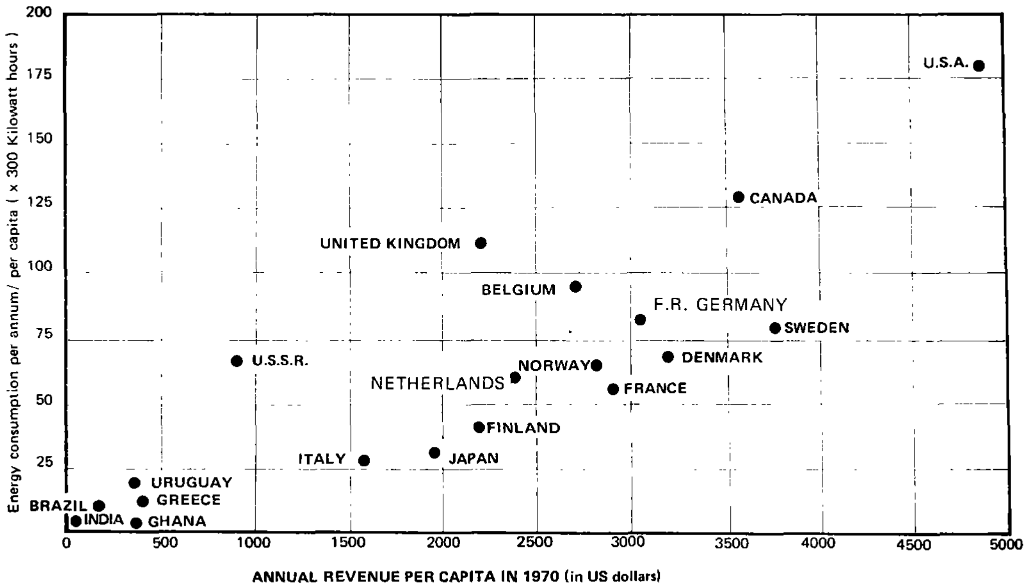


Fig. 43
Energy consumption and gross national product *per capita*. Expressed in litres of petrol consumed *per capita*/annum, this consumption is of the magnitude of 4,050 l in the U.S.A., 2,025 l in the United Kingdom and only 40,5 l in India (1970). Energy consumption and G.N.P. *per capita* present an appropriate correlation

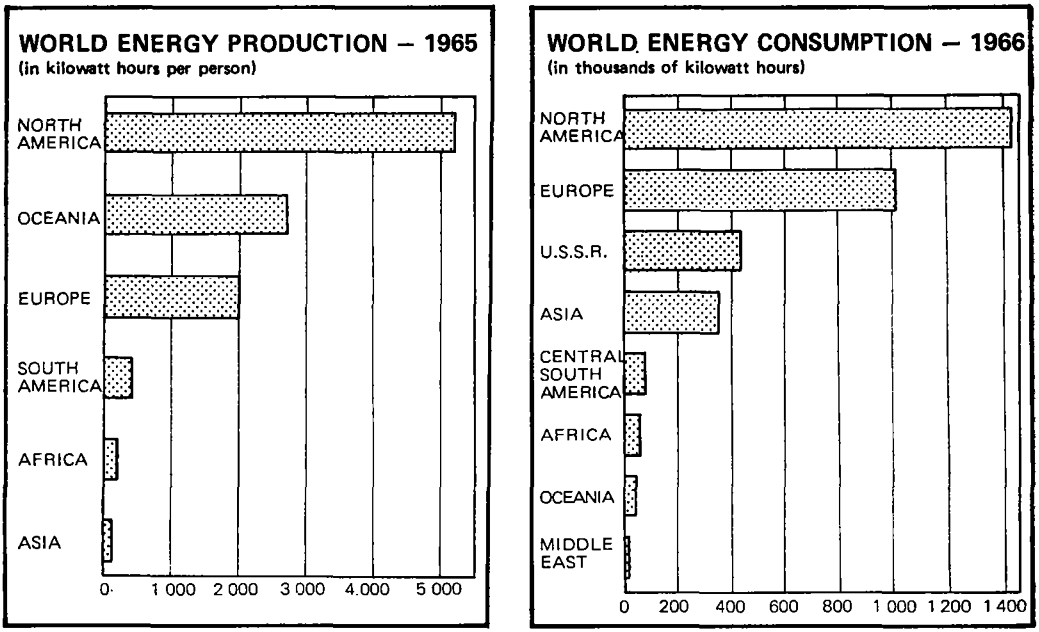


Fig. 44
Comparative histograms
of world energy production (1965)
and consumption (1966)
Source: J. Mchale. *World Facts and Trends.* London, Collier-Macmillan, 1973.

Table 12
Distribution of population
and production and consumption
of energy, 1970 (in percentage)

World and regions	Population	Production	Consumption
World	100	100	100
Africa	9.5	6.5	1.6
America			
North	6.3	32.7	36.2
Central } South }	7.8	5.1	2.1
Asia			
Middle East	2.1	13.1	1.1
Asia (n.e.s.)[1]	21.0	0.6	6.2
Asia (rest)	33.4	3.4	7.7
Europe			
Western	9.9	8.0	19.5
Eastern	3.0	5.8	6.5
U.S.S.R.	6.7	17.5	15.8
Oceania	0.5	1.0	1.1

1. People's Republic of China only.

Source: Calculated from the data given in the *U.N. Statistical Year Book, 1972.*

Table 12 lists world population along with energy production and consumption, by regions, for 1970. The United States, with only 6 per cent of the world's population, produces about 32 and consumes about 36 per cent of the total world energy. Japan (not shown in the table) with 3 per cent of world's population produces 0.8 and consumes about 5 per cent of the world's energy. The combined industrialized regions of the world (e.g. North America, Europe, the U.S.S.R., Oceania and Japan) constitute 30 per cent of the world's population but consumes 84 per cent of total world energy as against only 66 per cent production.

On the other hand, the Middle East with about 2 per cent of the world's population, produces 13 per cent of its energy supply. Since the region is not yet industrialized, its consumption of energy is very low, the surplus being exported to other industrialized countries like the U.S.A. and Japan with deficit energy budget (Fig. 41). In eastern Europe, the U.S.S.R. and Asia (mainly China) where economies are centrally planned, production almost matches with consumption.

One of the obvious reasons given for increasing energy consumption in the developed world is an increasing population demanding more and more goods and services. However, population increase does not explain the whole issue. For example, in the United States, the population increased by 43 per cent during the period 1946–66 while its energy consumption more than doubled. By the year 2000, the U.S. population is projected to be near 270 million, roughly a 30 per cent increase over the 1970 figure. Energy consumption, on the other hand, is predictably four or five times greater than at present. It appears, therefore, that the basic reason for increased energy demands is a changing way of life and a more intricate and expanding technology, rather than an enhanced population.

World energy prospects

Industrial growth in the world today is totally dependent on fossil fuels. Since their quantity in the earth's crust is limited and their rate of geologic production and accumulation immeasurably low in comparison with their rate of consumption, the ultimate exhaustion of this energy fund seems inescapable. Industrial man is, thus, living on a natural resources 'capital' rather than on an income. His initial 'capital' was the amount of energy stored up chemically from solar radiation over the last 600 million years. For how long can it sustain the industrial civilization? This is difficult to predict, since it involves projection of two independent curves of rates of production and consumption.

Table 13 shows estimates of initial reserves and those already drawn for consumption. The data for coal is by Paul Averitt (U.S. Geological Survey) and that for petroleum is by W. P. Ryman (Standard Oil Company, New Jersey).

The total energy content of the initial deposits of fossil fuels was about 64,000 million million thermal kilowatt hours. (See conversion factors below.)

Table 13
Fossil fuels: initial reserve and total withdrawal

Fuel	Initial reserve	Percentage of total energy content	Already withdrawn: From the beginning to end of 1970	Already withdrawn: In 1970
Coal	$7\,640 \times 10^9$ metric tons	89.8	140×10^9 metric tons	$2\,397 \times 10^9$ metric tons
Petroleum	$2\,100 \times 10^9$ barrels	5.5	268×10^9 barrels	17.96×10^9 barrels
Natural gas	$2\,000 \times 10^{12}$ cubic feet	4.7	276.9[1] $\times 10^{12}$ cubic feet	38.5×10^9 cubic feet

1. Production from 1951 to 1970.

Sources: M. King Hubbert. 'The Energy Resources of the Earth', *Scientific American* (New York) Vol. 224 (3), September 1971, p. 65; U.N. Statistical Year Books, 1950, 1960, 1972.

Conversion factors of sources of energy

Coal equivalent (metric tons) of
- One metric ton of crude petroleum (sp. gr.: 0.85) =1.3
- 1,000 cubic metres of natural gas =1.33
- 1,000 kwh of electrical energy =0.125

One metric ton of petroleum (sq. gr.: 0.85)=7.775 barrels
One cubic foot =0.02826592 cubic metres.

From the beginnings of human exploitation to the end of 1970 about 1,600 million million kilowatt-hours energy have been consumed. The remaining fossil fuels are made up of approximately 95 per cent of coal, 4 per cent of petroleum and 1 per cent of natural gas. The world petroleum consumption is currently rising at an annual rate of 6 per cent. Oil and gas now account for about 65 per cent of the world energy consumption and are expected to increase their relative position.

At the rate of a 60 per cent increase in energy consumption per decade, about 0.45 million megatons of coal equivalent of fossil fuels will be consumed between 1971 and A.D. 2000. Assuming a constant rate of energy consumption after A.D. 2000, the remaining stock of fossil fuels can be exploited up to A.D. 2330. If the fate of individual fuels are considered separately, then at the current rate of increase in demand, petroleum will be exhausted by the end of this century and natural gas a decade earlier. But in view of the fact that the rate of production of the mineable commodities pass through three stages, a period of ascent, a peak period, and a period of descent, it is expected that by 1980, or so, oil production will reach its peak and then start declining. A similar fate also awaits natural gas. It is suggested that coal will become much more valuable. In any event, all the fossil fuels will be rare commodities sooner or later; and the effect on the world economy has already been demonstrated by the recent restriction in oil availability by the Middle East countries.

Future alternatives

The crucial position of fossil fuels calls for a new orientation of the present set-up of industrial society towards other sources of energy. World fossil fuel supplies and trends are summarized in Fig. 45. Let us examine the alternatives.

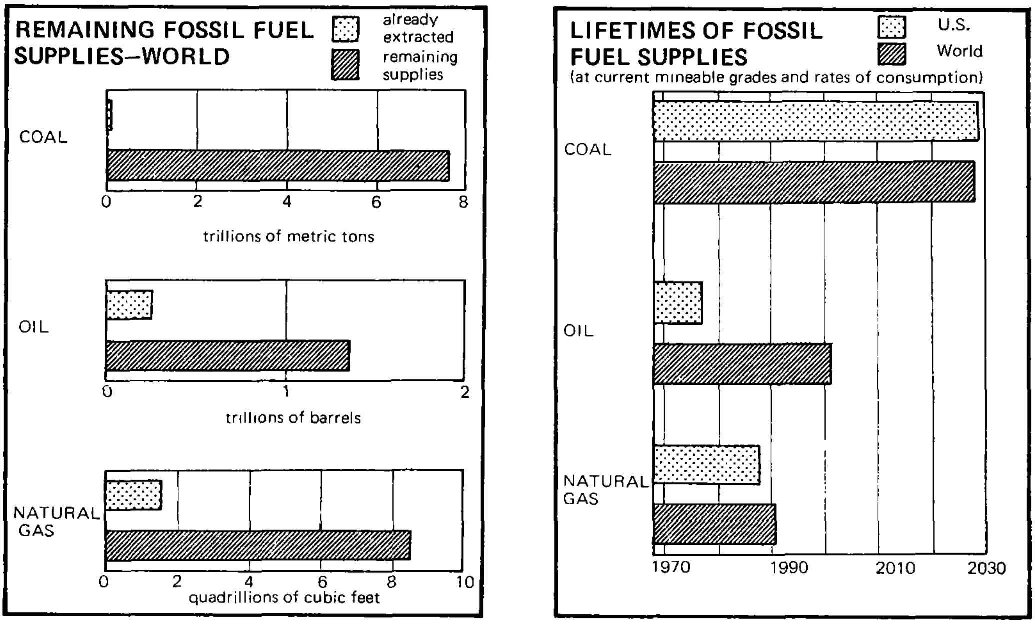


Fig. 45
Fossil fuel supplies and trends
Source: J. Mchale. *World Facts and Trends*. London, Collier-Macmillan, 1973.

At present, water power is widely in use all over the world. In 1960, the U.S.A. produced 149,515 million kwh of hydro-electricity; Japan 67,957 million kwh; U.S.S.R. 50,913 million kwh; India 7,874 kwh; U.K. 3,133 million kwh and Czechoslovakia 2,495 million kwh. Tidal power is exploited today in France, with a tidal electric plant on the channel coast which has a capacity of 240 megawatts. Geothermal plants occur in Italy (390 megawatts), California (400 megawatts) and New Zealand (290 megawatts). These exploit geothermal heat of volcanic origin.

Assuming full exploitation of all the suitable sites throughout the globe, estimates put maximum capacity of wind power to be in the order of 0.1 million megawatts; that of water power 3 million megawatts; that of tidal power 0.06 to 1 million megawatts; and that of geothermal power 0.06 million megawatts. But the total power capacity is not sufficient to meet future requirements of the world. Taking into consideration population increase, simultaneously, one would have to triple the world's production of man-made energy now in order to bring all peoples up to the minimum energy-use standard in advanced countries. In addition, the following problems cause a great concern for consuming nations, all over the world: increasing dependence on

a few sources of supply, mainly in the Middle East, to meet rapidly rising demands;[1] sizeable cost increase imposed by key producing countries; the increasing cost of finding and developing new energy sources; the many new ecological regulations, though meeting social objectives, nevertheless add significant cost burdens to the operation of the petroleum industry.

To overcome the critical situation, development and expansion of sources other than fossil fuels are indicated. Of such sources, only nuclear power seems to hold immediate promise at present, but wind and tidal sources, as well as the hydrolysis of sea water (see below), need serious consideration.

Nuclear energy as a resource

Nuclear energy is obtained either by a 'fission' or by 'fusion'. In atomic fission, the nucleus of uranium-235 is split; a fission of one gramme of this substance releases an amount of energy equivalent to the combustion of about 3 metric tons of coal or about 14 barrels of crude oil, but this process involves a rare isotope of a very rare element forming only 0.00018 per cent by weight of the earth's crust. World reserves of uranium are shown in Fig. 46. These may soon be exhausted but hope lies in the 'breeding' process in which uranium or plutonium is burnt in the 'breeder-reactor' in such a way as to convert thorium-232 or uranium-238 into fissionable material. Thus in the 'breeding process', more nuclear fuel is produced than consumed. A breeder-reactor may be expected to double the original quantity of fuel within six to twenty years time.

Nuclear fusion depends upon fusing isotopes of hydrogen bringing about a conversion into helium. The fuel for this purpose is deuterium which can be separated easily from sea water. Consumption of one atom of this fuel releases about 8 million million joules. It is estimated that the deuterium present in only 27 cubic kilometres (km^3) of sea water, can produce energy equivalent to the entire initial reserves of all fossil fuels. Another desirable feature of fusion power is that it produces no radioactive ash.

According to Dr. B. I. Spinrad (formerly with IAEA) 'we are, right now, at the threshold of a period in which the standard electric power station will be nuclear'. He quotes three reasons for this : the intrinsic cost of nuclear energy (in terms of the energy available from natural uranium) is one tenth to one-fifth of that for competitive fossil fuels; the intrinsic compactness of nuclear fuel makes it easily portable and reduces transport costs of fuel; the effluent from nuclear energy (fission products) is as compact as the fuel, and the environmental hazards due to the emission of nuclear energy effluent is thereby minimized.

On the other hand nuclear energy development is accompanied by a host of technical difficulties. Of more far-reaching significance is the environmental problem of disposal of radioactive waste from nuclear plants. An accidental release may contaminate all water circulation systems in the hydrosphere.

1. More recent developments, after 1972, towards the exploitation of oil in the Niger Delta (Nigeria) and of the sub-marine deposits of oils and gas around the United Kingdom, has shifted the limitation of dependence from the Middle East to a certain extent.

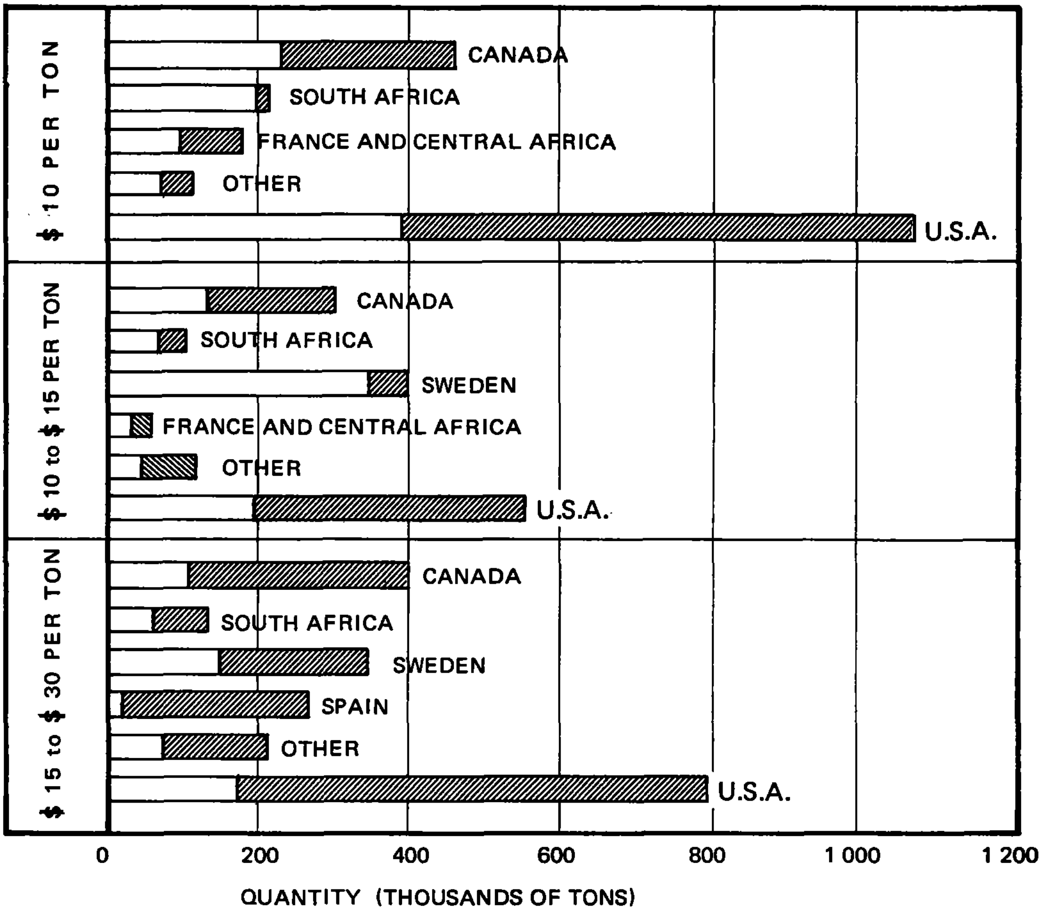


Fig. 46
World reserves of uranium given in tons of uranium oxide (U_3O_8). The uncoloured part of each bar represents reasonably assured supplies and the coloured part estimated additional supplies

Source: M. K. Hubbert. 'Energy Resources of the Earth', *Scientific American* (New York), September 1971.

Because of the harmful effects of any type of energy consumption, it is often suggested that the over-all use of energy must be arbitrarily limited. This, according to many, is equivalent to demanding the limiting of water supplies or food production. Many of the underdeveloped countries, at present, do not have available resources of energy to move into an industrial or advanced agricultural society. If the humane objective of providing a better quality of life for the inhabitants of these countries is maintained, then the essential role of energy in world affairs must be acknowledged.

5 Atmosphere

Introduction

The earth's atmosphere envelops the planet, in the same way as the atmosphere of other planets in the solar system, except that of Mercury, but its composition is unique, as far as our present knowledge stands. The atmosphere is, par excellence, very much dependent upon a balance of natural forces which sustain it. All of these, particularly gravitation, atmospheric pressure, the various gases, water vapour and various energy forces, have a profound effect upon our continued existence.

The surface pressure of the atmosphere is an index of the total mass of the envelope from the surface of the earth to its spatial outer limit; on earth, the average surface pressure is 1,015 millibars (1 mb = 1,000 dynes/sq. cm). Studies in the disciplines of aeronautics, atmospheric physics, geomagnetism, radio waves and satellite projects, have both confirmed a number of earlier interpretations of atmospherical structure and revealed much new data.

The earth's atmosphere is made up of a number of zones or layers, each with its own characteristics. There are boundaries between the zones which still require further study. Nearest to the earth's surface is the *troposphere*, containing the clouds and in which weather patterns are affected. Most of the water vapour in the whole atmosphere is found in this zone. The outer extent or boundary of the troposphere is from eight to twelve kilometres at high or middle latitudes and from sixteen to seventeen kilometres at the equator.

The *stratosphere* surrounds the troposphere and its chief characteristic is its constant temperature, i.e. it is isothermal. For reasons not fully determined, the stratosphere is generally absent in the tropic zone. A layer characterized by ozone, in which temperatures increase with outward distance up to 50 km, then progressively decrease up to 80 km, is called the *mesosphere*. Beyond the mesosphere is the *ionosphere* where the temperatures reach very high values as elevation increases towards the outer limits of the atmosphere. Table 14 shows the relationship between temperatures and pressures in the earth's atmosphere.

If we proceed to examine the nature of near-earth space, we observe a zone that is traversed by a reticulum of magnetic field lines, linked to the earth. The zone

Table 14
Temperature and pressure data in the earth's atmosphere

Altitude (km)	Absolute temperature (°K)	Pressure (mb)
0 (sea level)	288.16	1.013×10^3
11.02	216.66	2.263×10^2
20	216.66	5.475×10
25	216.66	2.489×10
32	237.66	8.678
47	282.66	1.204
53	282.66	5.832×10^{-1}
91	196.86	1.815×10^{-3}
128	273.6	1.451×10^{-5}
314	973.5	1.447×10^{-8}

also contains ionized gas that predominates over the neutral atmosphere. The magnetosphere represents the extreme, outer limits of man's environment; it has an effusion of ions and electrons derived from the earth's upper atmosphere and a plasma derived from the impinging solar wind. Scientists have been puzzled for many years about the cause and effect of magnetospherical phenomena; nevertheless such phenomena have been known to man for many centuries and described by classical writers. The famous auroras (e.g. northern lights) of the polar regions and magnetic storms feature in the behaviour of the magnetosphere. The magnetosphere protects our atmosphere from a direct interference by the solar wind and it shields the stratosphere (at low and middle latitudes) from what would be catastrophic doses of proton fluxes that are emitted by the intense solar flares. The magnetosphere poses problems for astronautical programmes, in so far as its radiation belt makes it unsafe to man satellites for lengthy periods. It also interferes, at times, with man's modern communication systems, such as short-wave radio transmissions and storms, associated with magnetic-field variations on the earth's surface, can be adverse and even put out of action the high tension electric power lines, on which our domestic and industrial supplies of energy depend. Even the rotation of the planet earth can be affected by magnetospheric disturbance. Today, the ICSU has inaugurated a programme of co-ordinated study of the magnetosphere and the years 1976–8 have been declared for the investigation, with a fundamental objective of exploring the extent to which man may exert control over the near-earth space that marks the periphery of his environment.

The ionosphere is much affected by the short-wave radiations from the sun. These radiations bring about a molecular dissociation of the gaseous constituents of the air, simultaneously becoming ionized by the liberation of free electrons. Thus, above 110 km this outer zone of the atmosphere forms an electrical conduction layer. Man makes particular use of the ionosphere in his applications of the principles of radio communication.

Composition of air

The gaseous components of the earth's atmosphere are known as air. Air is a mixture of gases but its composition is not constant, the variability in composition being affected by the interplay of a number of physical and biological factors. The gases which profoundly affect living systems are oxygen, carbon dioxide and nitrogen; because of the intimate relationship between these gases and the biosphere their properties will be discussed in more detail later in this chapter. Naturally occurring components of the air, variable in concentration, are water vapour and carbon dioxide. Oxygen, nitrogen, hydrogen, methane and nitrous oxide are also normal components of air and these are essentially non-variable in concentration. Substances such as formaldehyde, sulphur dioxide, ammonia and carbon monoxide are found in air, but these arise generally from man-made industrial operations. They are thus considered rather as contaminants than true constituents of air.

A general analysis of the air in the lower atmosphere is as shown in Table 15.

Table 15
Normal composition of the lower atmosphere

Gaseous component	Chemical formula	Volume (percentage)
Nitrogen	N_2	78.084
Oxygen	O_2	20.946
Argon	Ar	0.934
Carbon dioxide	CO_2	0.033
Neon	Ne	18.18
Helium	He	5.24
Krypton	Kr	1.14
Xenon	Xe	0.087
Hydrogen	H_2	0.500
Methane	CH_4	2.001
Nitrous oxide	N_2O	0.500

When we consider water vapour as a variable constituent of air then its variability in concentration is dependent upon the radiant energy of the sun (as heat energy) bringing about a conversion of water into water vapour, together with the attendant meteorological factors involved in mass movement of air, cloud formation, thunderstorm activity and heat-energy exchange. The amount of water vapour in air, at sea level, for example, may vary from 20 g/kg in the tropics to less than 0.5 g/kg in cold, dry, polar regions, and desert areas.

Other variable constituents of air, that should be considered in relation to our environment, are summarized in Table 16.

These substances are often confined to lower atmospheric layers in polar regions because of stratification of intensely cold air, but occur at considerable altitudes in equatorial and middle latitudes due to a vertical mixing of the air and general convection (winds). Common salt from sea water is the most common of these impurities. As population densities and industrialization increase, numerous other impurities also become prominent. (See the section on air pollution later in this chapter.)

Table 16
Variable constituents of the air of natural origin (excluding water vapour)

Meteoric dust (from other parts of the universe)	
Sodium chloride (chiefly from seas and oceans; salt lakes)	
Air-born soil (wind storms over dry land areas)	
Nitrogen dioxide (lightning flashes, electrical discharges)	
Sulphur dioxide	of volcanic origin
Hydrogen chloride	of volcanic origin
Hydrogen fluoride	of volcanic origin
Hydrogen sulphide (sulphur bacteria; volcanic origin)	
Ozone (by photochemical action; electrical discharges)	
Pollen (from all forms of phanerogams, chiefly grasses)	
Fungal spores	
Bacteria	
Viruses	

Sources: P. L. Magill; J. R. Holden; C. Ackley. *Air Pollution Handbook*. New York, 1956.
Sidney Licht. *Medical Climatology*. Baltimore, Waverly Press, 1964, p. 78.

Oxygen cycle

Oxygen occurs in the atmosphere as molecular oxygen (O_2) alternatively combined as ozone (O_3), carbon dioxide (CO_2) and water vapour (H_2O). The role of ozone in removing considerable ultra-violet radiation from solar radiation in the upper atmosphere has been mentioned in an earlier chapter. The major forms in which oxygen occurs are part of an intricately interrelated cycle of events involving organisms and numerous inorganic substances. Some of the roles of oxygen in these events are discussed below.

Oxygen is a major component of the three classes of foods (proteins, carbohydrates and lipids) and indeed of all living matter itself, comprising about one-fourth of its aggregate atoms. Since the fundamental route of inorganic matter to organic matter is through photosynthesis, it is appropriate to note again the summary statement of this process:

$$6CO_2+12H_2O \xrightarrow{\text{Radiation energy}} C_6H_{12}O_6+6O_2+6H_2O.$$

Through the use of a heavy isotope of oxygen, it has been shown that molecular oxygen (O_2) released during photosynthesis comes from the photolysis of water. This in turn suggests that the oxygen of the organic molecules synthesized, comes from the carbon dioxide. Thus

$$nCO_2+nH_2O^* \rightarrow (CH_2O)_n+nO^*+nH_2O.$$

(The (*) on the oxygen indicates a labelled isotope of the element.)

Molecular oxygen, as produced in photosynthesis, reacts readily with the organic compounds and other reduced substances. Oxygen thus serves as a hydrogen acceptor in aerobic respiration as summarized in the following statement:

$$C_6H_{12}O_6+6O_2 \rightarrow 6CO_2+6H_2O+\text{energy}.$$

The oxygen of the organic molecule appears in the carbon dioxide while the molecular oxygen serves as the hydrogen acceptor. These relationships between molecular oxygen, water, carbon dioxide and organic molecules can be summarized in simplified form as shown in Fig. 47. The respiration which converts the molecular oxygen back to water by reduction is accomplished by producers, consumers and decomposers.

Fig. 47
Simplified oxygen cycle

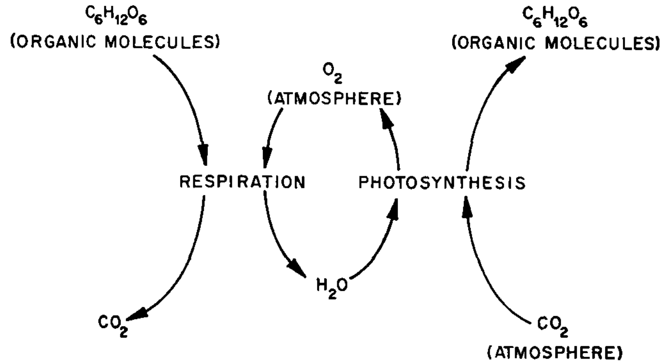
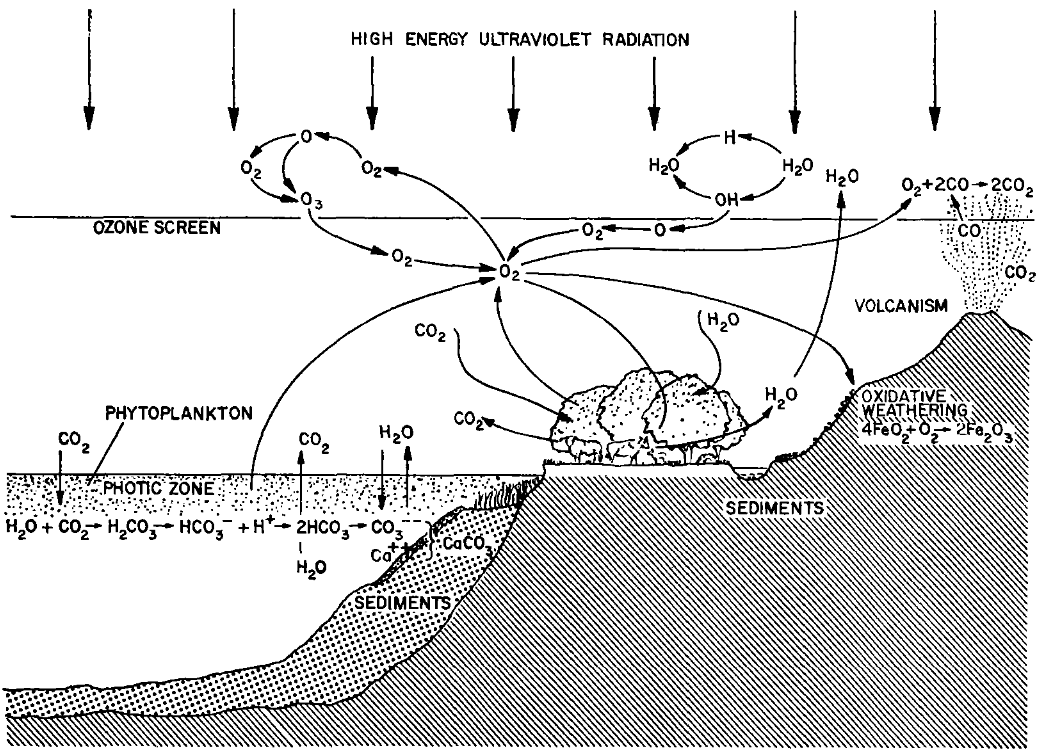


Fig. 48
Oxygen cycle
Source: Preston Cloud; Aharon Gibor. 'The Oxygen Cycle', *Scientific American* (New York), September 1970, p. 7.



The method of maintenance of the supply of these various forms of oxygen varies with the form. The principal reservoir of molecular oxygen is the atmosphere. Water resources are maintened in a cycle which will be discussed below in some detail. The carbon dioxide reservoir is constituted mainly in the atmosphere and the ocean waters. Maintenance of this reservoir of carbon dioxide is accomplished through respiration of organic food materials, of course, but also through such non-biological processes as volcanic eruptions, the burning of wood and the combustion of fossil fuels. The concentration of carbon dioxide in the atmosphere is maintained in dynamic equilibrium with that in solution in the ocean. Ocean waters are slightly alkaline and contain bicarbonate in a concentration of about 0.002 molar in equilibrium with atmospheric carbon dioxide. This equilibrium as well as several other major global pathways of oxygen are illustrated in Fig. 48.

The rates at which water vapour, molecular oxygen and carbon dioxide are processed by the biosphere vary greatly. As a result, the rates of exchange of the substances between the atmosphere, hydrosphere and lithosphere through processing in the biosphere vary among the three substances. Atmospheric carbon dioxide from respiration of animal and plant cells is utilized in a period of about 300 years and thus must be resupplied every 300 years. Atmospheric oxygen, with a much greater volume (21 per cent vs. 0.03 per cent), is used up about every 2,000 years. The earth's water is utilized in photosynthesis and regenerated in plant and animal respiration in an estimated period of about 2 million years.

Carbon cycle

Closely related to the events of the oxygen cycle and inseparable from them are the events and relationships of the carbon cycle. In fact, a comparison of Figs. 47 and 49 shows that both cycles revolve around respiratory and photosynthetic processes. The oceanic and atmospheric carbon dioxide (identical in the two cycles) is processed in photosynthesis to yield the organic carbon of the tissues of plants and animals (identical

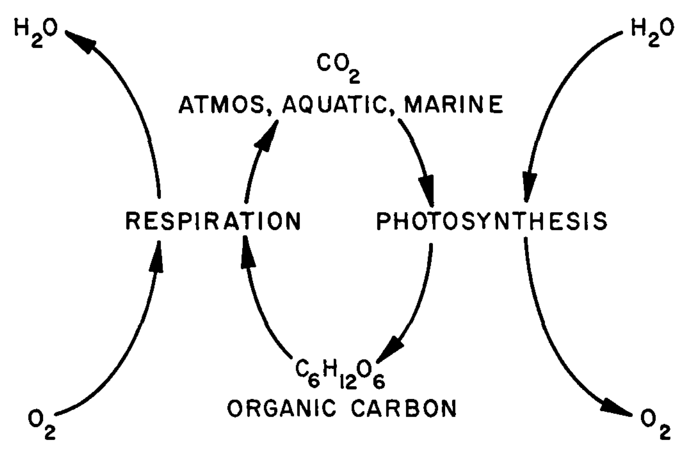


Fig. 49
Simplified carbon cycle

in both cycles). The major difference is that we are here tracing the cycle of carbon and thus do not need to concern ourselves with water or molecular oxygen.

The present level of carbon dioxide (1970) in the lower atmosphere is 322 ppm or 0.0322; this level has been augmented by approximately 10 per cent during the last century and it will probably increase again by 18 per cent towards the year 2000. An indication of the enhancement of carbon dioxide in the earth's atmosphere is illustrated in Fig. 50.

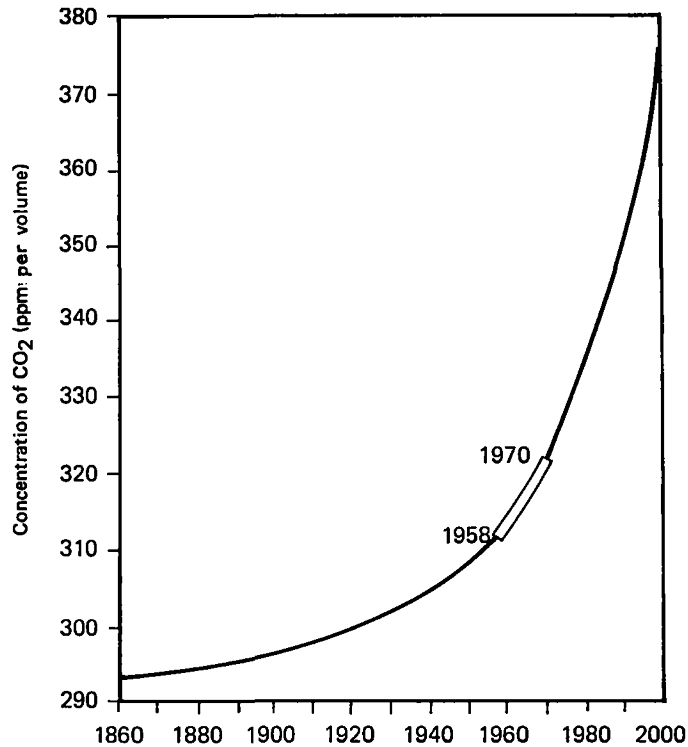


Fig. 50
The probable trend in increasing concentration of carbon dioxide in the atmosphere
Source: A. Sasson. *Développement et Environnement*. Paris and The Hague, Mouton, 1974.

Several significant aspects of the carbon cycle have not been included in order to simplify the analysis. First, in living communities, organic carbon is passed through the several trophic levels of food webs accompanied by digestion, transportation, and re-synthesis in the bodies of the organisms involved. Such changes are not supplied by Fig. 50.

Second, much is implied in the term 'respiration'. Producers and consumers respire, but the most substantial portion of the return of organic carbon through respiration is by decomposers. A very substantial return is accomplished also through activities often such as in the burning of wood as well as of coal, oil and gas which are the result of the transformation in geologic time, of decaying organic matter.

Third, the deposition of carbon takes many forms other than from the organic molecules of living tissues. Deposition of carbon in the organic form results when plant

constituents are highly resistant to microbial attack and accumulate as humus. Under certain conditions such as high moisture content, oxygen depletion and low pH, peat may accumulate. With the passage of time, compression of the deposits (along with other physical as well as chemical changes) results in the formation of coal. Another major form of carbon deposition in the organic form is found in mineral oil and natural gas. Inorganic deposition of carbon occurs when both carbonate and bicarbonate in the oceans (in equilibrium with atmospheric carbon dioxide) combine with calcium ions to form calcium carbonate ($CaCO_3$). This substance is also deposited in the shells of various animals (e.g. numerous molluscs, sponges, corals) and is released partially at least as a by-product of photosynthesis by some of the red algae in association with reef formation and also by certain green and blue-green algae. These sources of calcium carbonate provide an origin of limestone, a major constituent of the land mass of continents. These relationships have already been shown in Fig. 48 (oxygen cycle). The carbon of limestone eventually re-enters the cycle through weathering and the direct action of certain plants, e.g. algae, lichens. Solubility of calcareous deposits is increased by acid formation in such processes as nitrification and fermentation by micro-organisms.

Fourth, the interaction between atmospheric and oceanic carbon dioxide is of significance in the carbon cycle. These relationships are presented in the oxygen cycle and are not repeated here. Fig. 51 illustrates various relationships within the carbon cycle and the biosphere.

Some patterns of variation in atmospheric carbon dioxide concentration have been identified. The average concentration of carbon dioxide in the atmosphere is about 322 parts per million (ppm). If measurements are taken in a forest, minimum values of

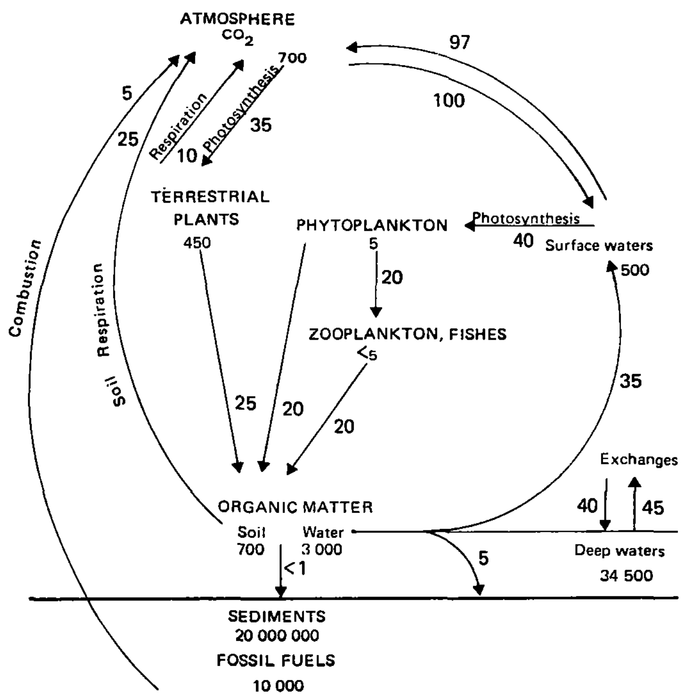
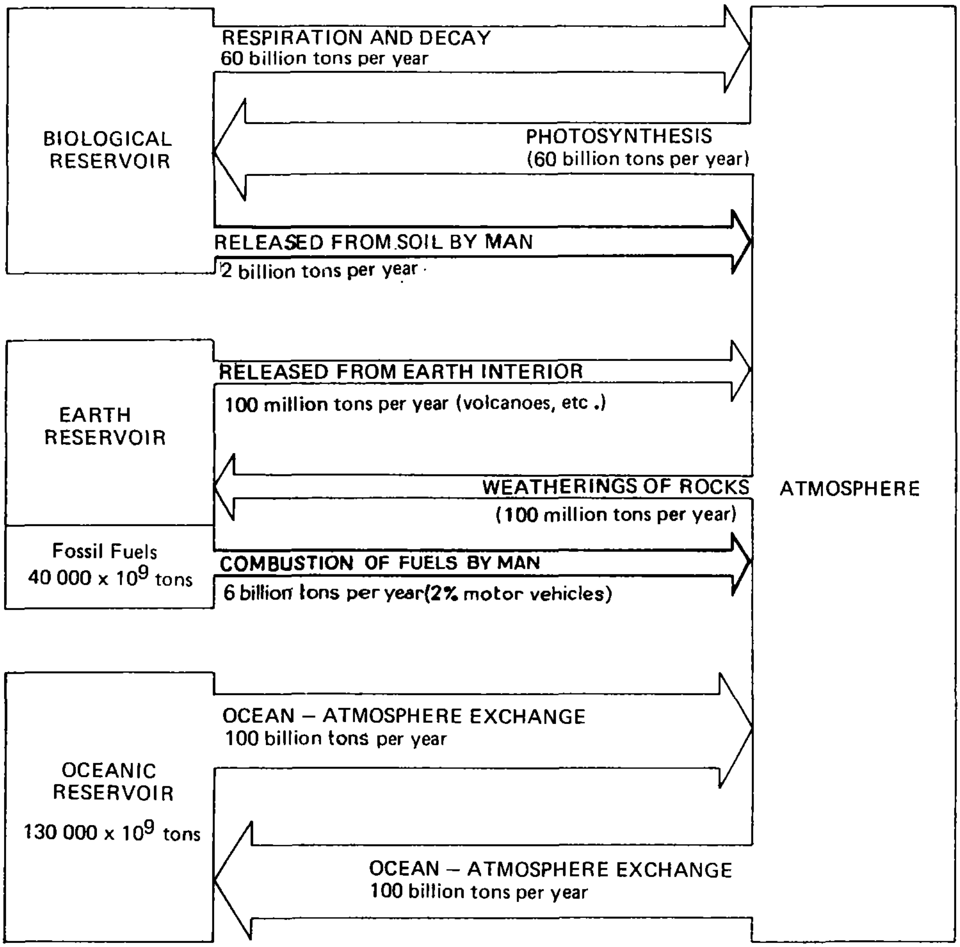


Fig. 51
The carbon cycle in the biosphere. (The large digits represent annual transfers of carbon as 10^9 tons in the biosphere; the small digits represent quantities expressed as 10^9 tons)
Source: A. Sasson. *Développement et Environnement*. Paris and The Hague, Mouton, 1974.



CARBON RELEASABLE FROM FOSSIL FUELS

RESOURCE	WORLD PRODUCTION IN 1967	REMAINING RESOURCES	CARBON ULTIMATELY RELEASABLE
Coal	115×10^9 tons	7270×10^9 tons	2250×10^9 tons
Petroleum	27×10^9 tons	1460×10^9 tons	244×10^9 tons
Natural gas	9.5×10^{12} m^3	864×10^{12} m^3	342×10^9 tons
Natural gas liquids	1.17×10^9 tons	121×10^9 tons	39×10^9 tons
Shale Oil	------	481×10^9 tons	370×10^9 tons

Fig. 52
The world's carbon dioxide balance. In this diagram one billion equals 1,000 million
Source: J. Mchale. *World Facts and Trends*. London, Collier-Macmillan, 1973.

about 10–15 ppm below the daily average will be found around noon at treetop level due to photosynthesis since sunrise. At around noon, temperature increases and humidity decreases have usually resulted in increased respiration rates so that the net consumption of carbon dioxide decreases. At night, photosynthesis stops, while respiration by producers, consumers and decomposers continues. As a result, the concentration of carbon dioxide increases and may exceed 400 ppm near the surface of the ground.

In some regions of the world, annual variations in carbon dioxide concentration of the atmosphere have also been identified. A marked decrease in concentration occurs during the spring as carbon dioxide withdrawal through photosynthesis greatly exceeds its return to the atmosphere from the soil.

Activities of man also influence the pattern of carbon dioxide in the atmosphere. As recently as the beginning of the 20th century, the over-all concentration of atmospheric carbon dioxide was 290 ppm, compared to 320–30 ppm in the early 1970s. This rapid increase is thought to be due primarily to the current domestic and industrial combustion of fossil fuels with release of carbon dioxide to the atmosphere. Based on fossil fuel consumption projections, it is estimated that the level may reach 375 to 400 ppm by the year 2000.

A very small proportion (a few tenths of a per cent) of the total bulk of carbon of the surface of the earth is in rapid circulation through photosynthesis, respiration and decomposition in the biosphere. As illustrated above, the changes in carbon dioxide concentrations that result from these three processes are measurable on a daily and seasonal basis demonstrating rapid changes in forms of carbon. The exchange between atmospheric and oceanic carbon dioxide also appears to be fairly rapid. Estimates based on use of radioactive carbon-14 indicate that a time interval of only five to ten years occurs from the time carbon dioxide is released into the atmosphere to its being dissolved in the sea. A summary of the world carbon dioxide balance is presented in Fig. 52.

Nitrogen cycle

Molecular nitrogen comprises about 78 per cent of the atmosphere and in this state it is of no use to the majority of organisms, yet the element is an essential component of living things. It ranks along with oxygen, hydrogen and carbon in importance. It is assimilated by green plants and becomes part of protein and nucleic acid molecules. As plant materials are eaten by animals, including man, these foods are digested and resynthesized to become parts of similar molecules in animal bodies.

However, atmosphere nitrogen cannot be assimilated directly by most green plants. Therefore, preliminary changes in the status of nitrogen must take place before the incorporation of nitrogen into living organisms. In fact, the route taken by nitrogen is quite complex, with many of the changes initiated by micro-organisms, especially the bacteria. Fig. 53 shows the major pathways and processes that are involved. The major processes are nitrogen fixation, nitrification, assimilation, ammonification and denitrification.

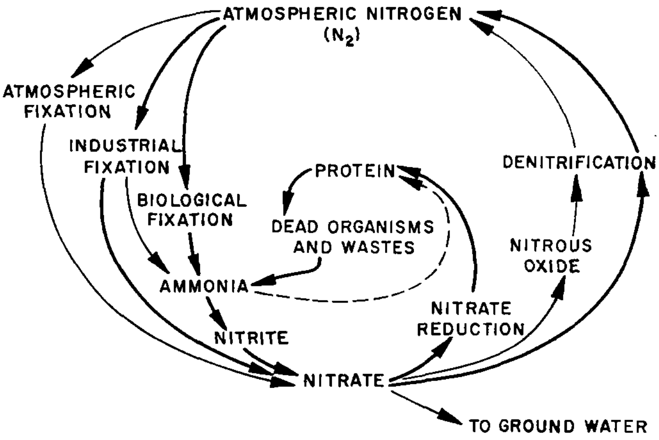


Fig. 53
The nitrogen cycle
Source: C. C. Delwiche. 'The Nitrogen Cycle', *Scientific American* (New York), September 1970, p. 4–5.

Nitrogen fixation

The over-all summary statement for the process of nitrogen fixation is as follows:

$$N_2 + 3H_2 \rightarrow 2NH_3.$$

The biochemistry of these reactions is incompletely understood, nevertheless ammonia is known to be the end-product of the fixation phase and the beginning of the nitrogen assimilation process. Such nitrogen is available in the atmosphere. Where the electrons originate is still a debatable issue, but current researches indicate that various sources of hydrogen are involved. By simulation, the nitrogen-fixation process can be performed industrially, the initiated technology for this being that of the German chemist, Haber; high temperatures and pressures are required to effect the chemical combination of nitrogen and hydrogen to produce the synthetic ammonia. By contrast in nature, nitrogen is combined chemically with either oxygen or hydrogen or both through cosmic radiational activity, meteor tracks and lightning.

Biological fixation is effected by various organisms that have the ability to reduce atmospheric nitrogen to ammonia. Such organisms include species of heterotrophs such as *Azotobacter vinelandii*, which is an aerobe and *Clostridium pasteurianum*, which is an anaerobe; the photosynthetic bacterium, *Rhodospirillum rubrum*, together with various forms of blue-green algae also contribute in nitrogen fixation. A special case of symbiosis, involving the bacterial genus, *Rhizobium*, together with certain phanerogams, chiefly Leguminosae, in which a biochemical liaison exists between the microbe and the higher plant, adds to the range of nitrogen-fixation systems known to man. Indeed, several million tons of nitrogen per annum are converted to ammonia under conditions far more subtle than in man-made systems.

The best studied systems involved in biological nitrogen fixation are those found in *Clostridium pasteurianum* as well as among the blue-green algae (Cyanophyceae) and the photosynthetic bacteria (Thiorhodaceae, Chlorobacteriaceae and Athiorhodaceae)

that fix both nitrogen and carbon dioxide. They play an important role in the cycle of nitrogen, most particularly in both nitrogen-deficient environments and in eutrophic waters. In tropical agriculture, both groups of these organisms play an indispensable part in raising soil quality.

Certain photosynthetic bacteria fix nitrogen only under anaerobic conditions; the enzymic complex known to function in this fixation process is appropriately termed nitrogenase. The complex is found equally among the blue-green algae. It is highly sensitive to the presence of oxygen, and yet these groups of organisms do absorb CO_2 and release oxygen in photosynthesis; how this happens is not fully understood, but it is suggested that there are intrinsic protective mechanisms, of morphological and physiological types, to be found in these intriguing plants.

Blue-green algae can survive with or without oxygen, although their supposed anaerobic status is still under investigation. They are highly adaptable organisms, withstanding extremes of temperature, dessication and other adverse environmental conditions. This successfully adjustable adaptation explains their ubiquity in many types of habitat, including eutrophic lakes, arid desert soils, thermal springs, seas and soils. They make the major contribution of combined nitrogen in nature. In the tropics, nitrogen fixation by blue-green algae may be 79 kilogrammes/hectare per annum; by comparison, at about 40° latitude S (the level of New Zealand), fixation by *Rhizobium sp.* has been estimated as high as 560 kilogrammes/hectare per annum.

Over-all and from an agricultural viewpoint, the bacterial symbiotic associations with legumes are of the greatest significance apropos human populations and productivity. Tropical legumes under favourable conditions may exceed the nitrogen-fixing capacity of their temperate counterparts. For example, nodular fixation in the pigeon pea added 39.46 kg of nitrogen/hectare and this amount was increased to 60.32 kg/hectare. Legume cover crops have been used very successfully in many tropical areas. *Glycine javanica* grows in Kenya soils, with appropriate *Rhizobium* symbionts, has added fixed-nitrogen, equivalent to 381 kg of ammonium sulphate per acre per annum.

An additional source of ammonia in nature is through the decay of dead organisms and the metabolisms of wastes through the process of ammonification. This process is discussed later in this chapter.

Nitrification

Most organisms cannot make use of ammonia directly in the conversion of nitrogen into protein and nucleic acids in living tissues. In nitrification, ammonia is oxidized (e.g. *Nitrosomonas* and *Nitrobacter*) to nitrites and to nitrates. These two processes are summarized as follows:

$$NH_3 + 1\frac{1}{2}O_2 \rightarrow HNO_2 + H_2O$$

$$KNO_2 + \frac{1}{2}O_2 \rightarrow KNO_3.$$

Each of the two steps provides energy for the bacteria involved. Thus the major form (nitrate) in which nitrogen is assimilated is available. Note, in review, that in nature this nitrogen has two major sources: nitrogen fixation from the atmosphere, ammonification using dead organisms and animal wastes.

In agricultural practice, the goal is to provide nitrate, so fertilization with manure depends on ammonification of organic nitrogen and subsequent nitrification of the ammonia. Fertilization with dilute solutions of ammonia depends on nitrification. Nitrates are also used in fertilizers, but they are highly soluble in water and tend to leach from the soil and be transported by water. Ammonia also contains a greater quantity of nitrogen per unit of weight than nitrate. There is now an extensive use of fixed ammonia as a fertilizer.

Nitrogen assimilation

Nitrates (and less significantly, ammonia) serve as the nitrogen sources for assimilation by most green plants. When nitrates are used, they are reduced in the living cell to ammonia. The ammonia is then incorporated into the amino groups ($R\text{-}NH_2$) of amino acids such as glutamic acid and aspartic acid. These two amino acids are the precursors of the various nitrogen-containing substances (mainly proteins and nucleic acids) in living things.

Ammonification

When plants and animals die, the organic nitrogen of their tissues is metabolized by micro-organisms and ammonia is produced. This process is called ammonification and is accomplished by various bacteria and fungi. Proteins and nucleic acids are first converted to amino acids and nitrogenous bases, and then removal of the amino groups (deamination) may take place. This process is summarized below, using the amino acid, glycine.

$$CH_2NH_2COOH + 1\frac{1}{2}O_2 \rightarrow 2CO_2 + H_2O + NH_3$$

Under anaerobic conditions, amines are formed and ammonia is not immediately produced.

Green plants do not excrete nitrogen-containing wastes, but animals do. Invertebrates commonly excrete nitrogen as ammonia, but vertebrates excrete urea and uric acid. These substances are also metabolized by micro-organisms. Using urea (carbamide) the reaction is as follows:

$$O{=}C\begin{matrix} \diagup NH_2 \\ \diagdown NH_2 \end{matrix} + H_2O \rightarrow CO_2 + 2NH_3$$

After ammonification, nitrification may again occur. The nitrates are then again used in the production of proteins and nucleic acids.

Denitrification

The natural process which results in the return of molecular nitrogen to the atmosphere is termed denitrification. It is performed by various bacteria and fungi. Several stages and pathways of the process are known, but the net effect is to return the molecular, gaseous nitrogen (N_2) to the atmosphere. The basic step in denitrification to molecular nitrogen occurs under anaerobic or partially anaerobic conditions.

Man and the cycles

Oxygen

Barometric pressure decreases with increasing altitude, so the partial pressure of oxygen decreases correspondingly. For non-residents of high altitudes, some cerebral effects due to hypoxia (a deficiency of oxygen in the body) often occur at around 3,650 m. At this altitude about 50 per cent of the population will experience illness, headache and insomnia. At 6,000 m impaired consciousness often takes place after about ten minutes.

Considerable acclimatization to high altitudes is possible as evidenced by experienced mountain climbers and by numerous permanent human populations living at high altitudes. An estimated 10 million people dwell permanently at altitudes of 3,650 to 4,000 m in the Andes of South America and in and around Tibet. In most of those high-altitude regions, human habitation is limited more by climate than effects of altitude. In Peru, Bolivia and Chile, mining communities, which depend on food brought from lower elevations, occur quite commonly at 4,900 to 5,200 m. One community is found at 5,300 m and the miners climb an additional 450 m to work each day. They attempted to occupy a camp at 5,600 m but abandoned it because of loss of weight, loss of appetite and inability to sleep. It has been concluded that an altitude of 5,300 m is a maximum altitude for most people and that many would be unable to adjust even at that height above sea level.

The main symptoms of high altitude health deterioration are lassitude, increasing lethargy, loss of weight and, in some cases, vomiting. In early mountain expeditions, there was also evidence of dehydration and starvation.

Carbon

The greatest proportion by far of the carbon of the surface of the earth occurs as inorganic deposits (mostly carbonates) and organic deposits (oil shale, coal and petroleum). At the present time, human populations are introducing a major modification in this normally slow aspect of the carbon cycle through the rapid domestic and industrial combustion of the organic deposits with the release of carbon dioxide to the atmosphere. It was suggested above that the magnitude of this influence is revealed in the increasing concentration of carbon dioxide in the atmosphere. The long range effect of this change on the biosphere and man are not known. Global climatic changes have been hypothesized, based on the absorption of infra-red radiation by

carbon dioxide, but actual occurrence of such climatic changes has not been demonstrated. This warming effect on climates may be partially or entirely offset by the cooling effect of particulate air pollution, currently being emitted. Clearly, more information is needed before this global 'experiment' is carried too far.

Nitrogen

It has been in the interest of human welfare and technological programmes to increase food production. Extensive industrial nitrogen fixation in fertilizer production and the cultivation of legumes have made a large contribution to such food production. However, there has not been a similar incentive to increase denitrification processes, so there is reason for concern that denitrification may not be keeping pace with nitrogen fixation. No specific data is available but more information on the relationship of these processes is needed, since a serious imbalance in the nitrogen cycle could occur. Such an imbalance would have world-wide implications; the nature of its effects on human populations is highly speculative at this time. Fertilizers are estimated to

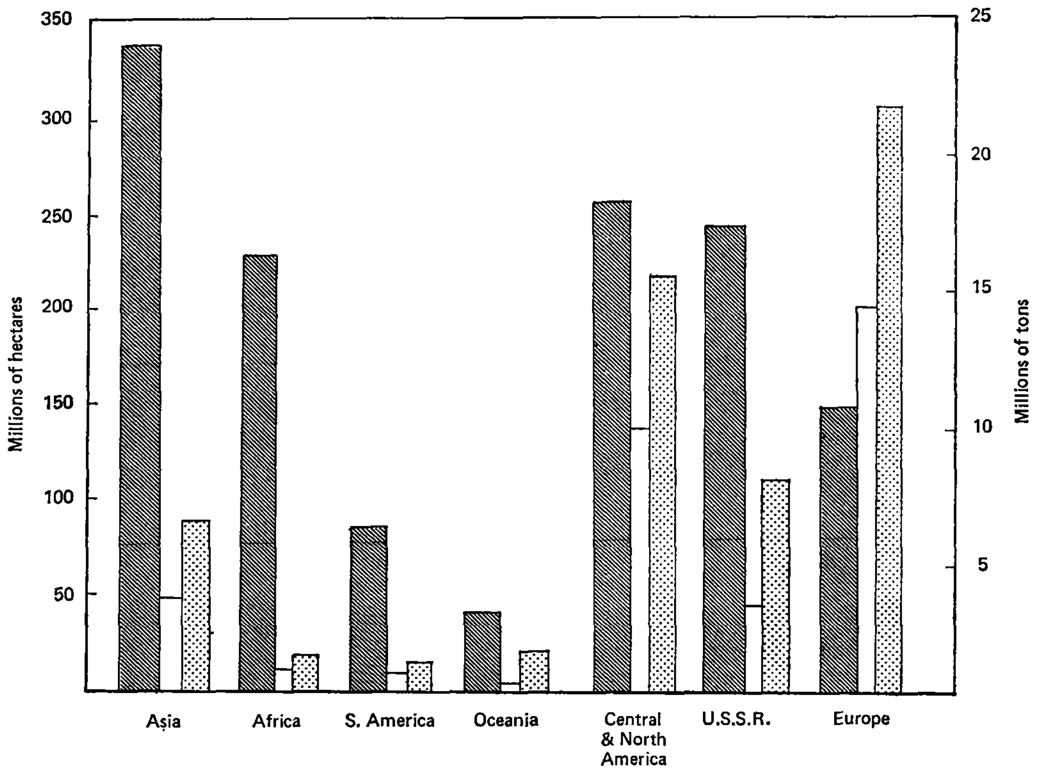


Fig. 54
Trends in the consumption of inorganic fertilizers in seven regions of the world
Source: FAO.

Cultivated areas: hatched
Chemical fertilizer consumption: 1962–63 white, 1967–68 dotted

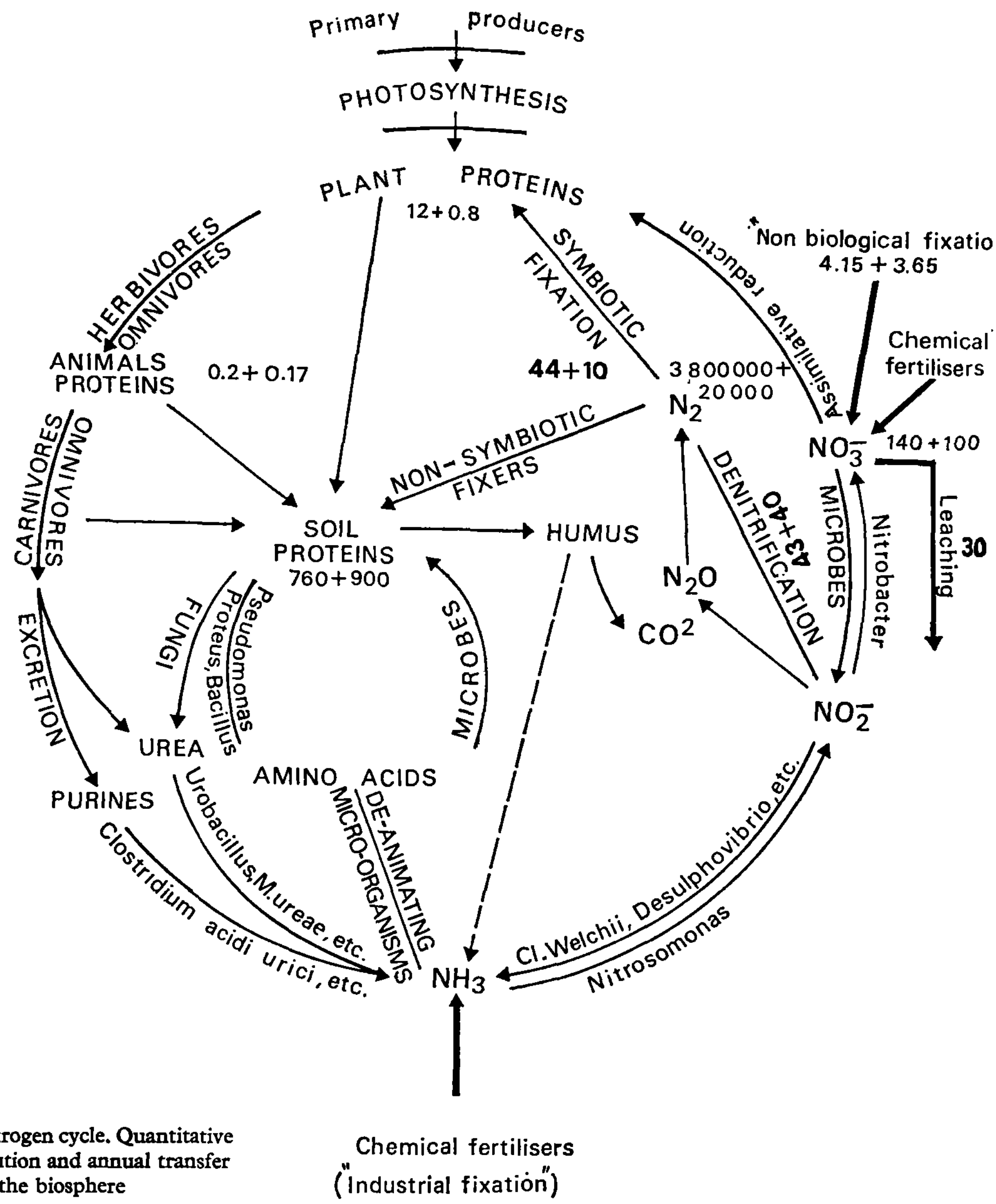


Fig. 55
The nitrogen cycle. Quantitative distribution and annual transfer within the biosphere

Source: A. Sasson. *Développement et Environnement*. Paris and The Hague, Mouton, 1974.

Note. Small digits represent nitrogen in gigatonnes. (The first number = quantity of N in the atmosphere and in living land forms; the second number = quantity of N in water and aquatic organisms.); large digits represent the quantities of N transferred each year in the biosphere. (The first number = transfer in the terrestrial sphere; the second number = transfer in the aquatic sphere.)

require a 30 per cent over-all increase to attain an adequacy in world food production and it must be recalled that almost 500 million people already rely upon heavily fertilized land, so the limits, on this basis, to population increase are already delineated. There are drastic side-effects of increased fertilizer production, e.g. one ton of manufactured nitrogen fertilizer, uses 5 tons of coal (or fuel equivalent). An illustration of chemical fertilizer consumption related to arable land is given in Fig. 54. (See also Fig. 55 for the nitrogen cycle.)

Atmosphere pollution

Most of our knowledge of the composition of air is based upon samples that have been taken very near the earth's surface. The composition of normal air has been well established, hence any components that are found in air samples that are over and above the normal analyses are to be considered as contaminants. Contamination of the atmosphere is not a recent event, nevertheless the intensity of air pollution and its implications in the process of survival are of current concern.

Although the atmosphere becomes contaminated on occasion by gases emitted during volcanic eruptions and related geological phenomena, this activity is not as significant as man-made pollution, resulting from commercial, industrial, agricultural and domestic operations. Man-made pollution of the environment, in general, is logically our own problem. The phenomenon is a major symptom of man's interference with nature in modern times; it shows itself specifically with an increasing concentration of carbon dioxide in the atmosphere, the mortality rate in man due to emphysemas, the frequency of lung cancers in workers exposed to sulphurous fumes and substances like perchloro-methylmercaptan.

The chief polluting agents of the atmosphere are: sulphur anhydrides (SO_2 and SO_3), dust particles associated with sulphurous emissions and often containing lead residues, carbon monoxide, photochemical oxidants arising from solar radiational effects on hydro-carbons and nitrogen oxides (NO, NO_2 and N_2O); other photochemical reactions, involving ozone, aldehydes and other chemical substances emitted into the atmosphere. A direct, adverse effect on the environment is the depletion of visibility and light quality by the presence of these contaminants in air; 'smog' results from such processes, being a water-vapour fog condensed around millions of smoke particles. Table 17 provides a classification of major air pollutants, detectable in many parts of the world, but chiefly in industrialized areas. (It can apply to Japanese towns, Hong Kong and Singapore, for example.)

The types of pollutants listed must be taken only as examples in general, since the impurities and their concentrations vary with such contributive factors as combustible fuels used in domestic usage, types of transport and the degree and types of industrialization. In the underdeveloped areas, i.e. areas where mechanizations and processing technology are absent, there are still many air pollutants to be found, arising from the burning of wood, charcoal making, bush fires, lime-kiln operations, etc.

By comparison, Table 18 presents a classification of air pollutants and their sources in an advanced, technological society.

Table 17
Classification of air pollutants and their sources

Pollutants	Sources
Smoke, grit and sulphurous gases	Incomplete burning of fuels such as oil and woodwaste in boilers, furnaces and incinerators. Burning of domestic and garden wastes in the yard
Smoke, carbon monoxide, nitrogen oxides, hydro-carbons and many other gases	Incomplete combustion in motor cars and diesel vehicles; also from aircraft engines
Smoke and other invisible pollutants	Ships in harbours and trains and airplanes during take-off
Dust	Industrial plants like steel works, cement works, ceramic factories and asphalt plants, construction sites and earth works
Smoke and sulphur dioxide	Oil refineries
Odorous gases	Processing and canning factories
Other undesirable and obnoxious gases	Chemical plants manufacturing a variety of chemicals like acids, sodium hydroxide, detergents, plastics, fertilizers and chlorine
Carcinogens	Benzopyrene and other related substances

Table 18
Air pollutants by source in a technological society[1]

Source	Millions of tons discharged yearly				
	Sulfur oxides	Nitrogen oxides	Organic compounds	Carbon monoxide	Particulates
Motor vehicles	1	6	12	66	1
Industry	9	2	4	2	6
Power plants	12	3	< 1	1	3
Domestic heating	3	1	1	2	1
Refuse disposal	< 1	< 1	1	1	1

1. Pulp and paper mills, iron and steel mills, petroleum refineries, smelters, fertilizer manufacturers, synthetic rubber manufacturers.

Source: P. Walton Purdom (ed.). *Environmental Health*. New York and London, Academic Press, 1971, p. 194.

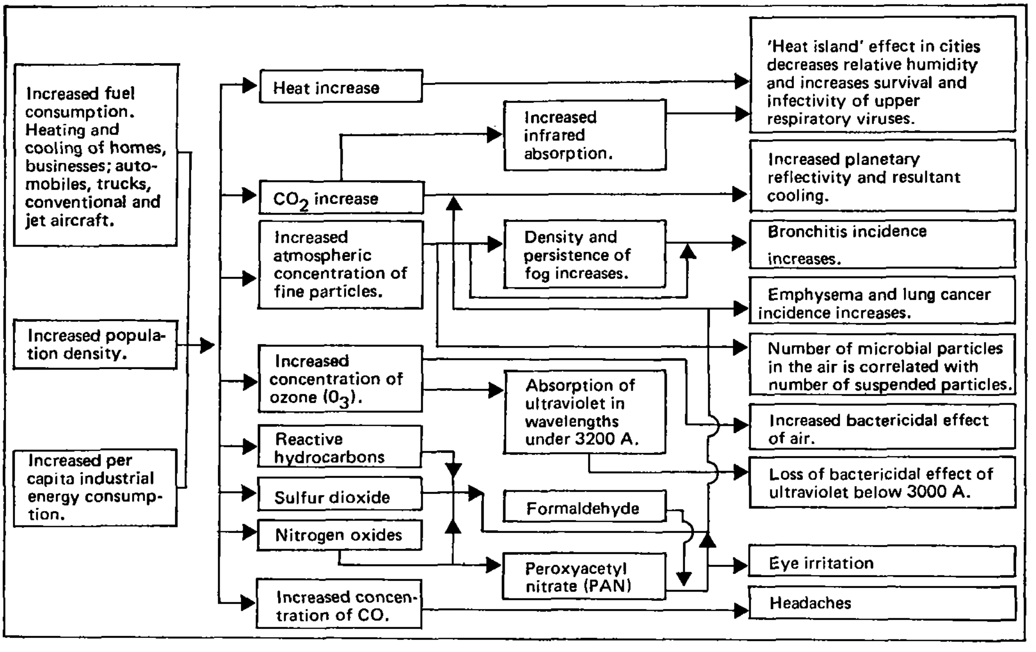
Weather and air pollution conditions

Weather exerts a strong influence on air pollution conditions. Strong winds effectively disperse pollutants over wide areas preventing a build-up to relatively high concentrations. However, if winds die down or a temperature inversion occurs over an area of land, then a stagnant air mass may develop. Atmospheric temperature decreases

with increased altitude, but in a temperature inversion, this pattern is reversed in an altitude range of approximately 600 metres beginning at an altitude of 900 metres, or so, and extending upwards. The effect of this condition is to prevent the normal rise, cooling, and dispersion of warm surface air and pollutants. Limited air dispersal allows the concentration of pollutants to build up as they continue to be emitted. Since weather conditions are also partly caused by local topography, the location of industrialized plants becomes important. Local weather and topography have been significant factors in numerous instances of build-up of air pollutant concentrations and resultant episodes involving human illness and death.

A second atmospheric condition influencing air pollution conditions involves the combination of sulphur dioxide fumes and fog to form sulphur dioxide smog. (Cf. earlier comment on particulate smog.) Combustion of coal results in emission of gaseous sulphur dioxide (SO_2) as a waste. A stagnant air mass in an area with high SO_2 emissions accompanied by fog thus often results in a heavy blanket of smog. Some air pollution episodes involving human illness and death have been caused by these specific conditions. Air pollution and human ecological effects are illustrated in Table 19.

Table 19
Air pollution and human ecological effects



Source: J. Mchale. *World Facts and Trends*. London, Collier-Macmillan, 1973.

Some effects of air pollutants

Sulphur dioxide as an air pollutant accelerates corrosion of metal, deterioration of certain fabrics and dyes, deterioration of leather and paper materials, erosion of marble, limestone and certain sandstone substances, and discolouration of lead-base paints. Ozone accelerates the cracking of rubber by attacking unsaturated carbon-carbon bonds in the rubber molecule. Sulphur dioxide, ozone, nitrogen oxides, and fluorides are some of the air pollutants known to affect vegetation. Symptoms of injury vary with the type of pollutant and the type of plant. Since certain plants are quite sensitive to certain air pollutants, they are often used as air pollution indicators. Fluorides from phosphate fertilizer production and aluminium (bauxite) refining affect dairy cattle when concentrations exceed 30–50 mg per litre in herbage. Symptoms include periodic diarrhoea, lameness, worn teeth and loss of body weight.

The adverse effects of man-made pollutants on people, as individuals or in communities, can be classified into three main categories, namely, disturbances in the respiratory tract; irritations of the eyes, nose and mouth parts, particularly the lips, tongue and pharyngeal areas; other ill-effects, for example, in the digestive system and of the skin. Consideration will be given here to respiratory tract disturbance, due to atmospheric pollution.

The major human hazard of air pollution is associated with the respiratory system. The main respiratory diseases associated with air pollution are chronic bronchitis, emphysema, bronchial asthma and lung cancer. In chronic bronchitis, inflammation of bronchial tubes, reduction in ciliary action, excessive production of mucus, and bronchial infection are often involved. A chronic cough and shortness of breath are typical symptoms.

Emphysema involves the abnormal inflation of the alveoli of the lungs, a condition which arises from the constriction of the bronchioles. Alveoli sometimes burst and adjacent ones united. Lung efficiency is thus reduced and shortness of breath becomes chronic.

Bronchial asthma involves an allergic reaction of the bronchial membranes to various substances. The membranes swell, causing an obstruction of the bronchioles and the resultant shortness of breath and breathing difficulty.

Lung cancer involves the appearance of malignant cells in the lining of the respiratory tract. Many factors including increased life span appear to be related to the current increasing incidence of this disease. There is considerable evidence that substances called carcinogens also stimulate lung cancer. Examples are benzopyrene found in coal smoke and cigarette tars inhaled by smokers.

Numerous studies of children and adults have shown evidence of additional less-pronounced health and performance conditions related to air pollution. These include decreased breathing capacity in children and a decreased physical performance by athletes.

Finally, numerous episodes of high air pollution are on record which have involved the illness and deaths of many people. These have generally occurred in high population density areas of technological societies. They have usually been precipitated

by a combination of high pollution emissions and unfavourable weather conditions.

Perhaps the influence of air pollution on human health is the most compelling one, but other considerations are also pertinent. Some losses in crop yield and animal health are inevitable and the general deterioration of various inert materials already cited earlier in this section, results in losses to the community.

Manufacturing, administrative legislature costs for environmental quality control must also be met. Aesthetic losses are often expressed as economic losses as well as losses in human mental and emotional well-being.

Interactions of the atmosphere and human populations: a summary

In this chapter, various interdependencies between the atmosphere, biosphere and inert materials, have been outlined. Several specific relations with man have also been mentioned. First, through rapid combustion of fossil fuels, carbon dioxide is being returned to the atmosphere more rapidly than it is being removed. This is currently resulting in a rapid change in the concentration of atmospheric carbon dioxide. Long-term effects of this change are not known. Second, nitrogen of the atmosphere is now being fixed industrially at a rate at least equivalent to total nitrogen fixation in the biosphere before modern agriculture. This results in extensive changes in the nitrogen cycle with unpredictable effects. It also results in high concentrations of nitrates in inappropriate places like human water supplies, which are hazardous to human health. Third, differing partial pressures of atmospheric oxygen and carbon dioxide have various effects on the human body and the human body appears to be capable of some adaptation to such variations. Fourth, various impurities are being added through technology which may have harmful effects on human health and unpredictable long-term effects on the biosphere as a whole. The problem of human ecosystem management, in the light of these and other interactions, is discussed in a later chapter.

6 Water

Introduction

The purpose of this chapter is to provide a brief analysis of the influences of water on human populations. The major considerations pertain to water quantities and the hydrologic cycle, several major relationships between water and organisms, world distribution of water, purity and impurity of water, natural purification and major human sources of water. The hydrologic cycle and relationships between water and organisms are basic to an understanding of water as an environmental factor vital to all forms of life. World distribution of water relates to the distribution and densities of human populations. The quality of water affects the health of man.

The hydrologic cycle

About 75 per cent of the earth's surface is covered with water. The oceans and seas comprise about 97 per cent of this and about 2.15 per cent occurs as ice in the polar ice caps and glaciers. The rest is made up primarily of underground water, surface water and atmospheric water vapour, in order of decreasing amounts. Of the total, only about 0.01 per cent occurs as atmospheric water vapour at any given time, amounting only to an average of 25 millimetres of rain over the entire surface of the earth. Table 20 provides estimates of the water component of the various parts of the earth. Atmospheric humidity levels rise near the equator and are equivalent to about 44 millimetres of precipitation. Around 40 to 50 degrees latitude, the mean annual yield is about 30 millimetres. There are large variations from these figures due to the influence of air currents and geographical features. Table 21 provides data on the water balance, expressed in flow-intensity units (volume/area × time) for the oceans and land masses.

The principal pathways of water throughout the major parts of the earth are collectively referred to as the hydrologic cycle. Fig. 56 presents the hydrologic cycle along with estimates of the volumes of water involved, in cubic kilometres per day. Many specific additions and modifications could be superimposed on the cycle. For

Table 20
Distribution of the world's water resources

Location	Surface area (sq. km)	Water volume (km^3)	Percentage of total
Surface water			
Freshwater lakes	860 000	125 000	0.009
Saline lakes and inland seas	700 000	104 000	0.008
Average in stream channels		1 250	0.0001
Sub-surface water	130 000 000		
Soil moisture		67 000	0.005
Groundwater within 0.8 km depth		4 170 000	0.31
Groundwater, deep-lying		4 170 000	0 31
Total liquid water in land areas		8 637 000	0.635
Ice caps and glaciers	18 000 000	29 200 000	2.15
Atmosphere	510 000 000	13 000	0.001
World oceans	360 000 000	1 322 000 000	97.2
TOTAL (rounded)		1 360 000 000	100

Source: Raymond L. Nace. *U.S. Geological Survey*. Washington, 1964.

Table 21
Water balance of oceans and lands (cm/year)

	Precipi-tation	Evap-oration	Run-off	Inflow (—) or Outflow (+)
Oceans				
Atlantic	89	124	—23	—12
Indian	117	132	— 8	— 7
Pacific	133	132	— 7	+ 8
World Ocean	114	126	—12	0
Lands				
Africa	69	43	26	
Asia	60	31	29	
Australia	47	42	5	
Europe	64	39	25	
North America	66	32	34	
South America	163	70	93	
Total land areas	73	42	31	

Source: M. I. Budyko. *Climate and Life*. Leningrad Hydrological Publ. Ho., 1971.

example, water falling on the land actually strikes vegetation, some also strikes the soil and is held there until it evaporates, while some is taken up by the roots of plants and is then lost to the atmosphere through transpiration. Water evaporates while falling as rain, and some evaporates from rivers and streams. All of the water which does not evaporate is accounted for by surface run-off, together with infiltration into and through the soil and also by percolation to underground water levels and even deep rock percolation.

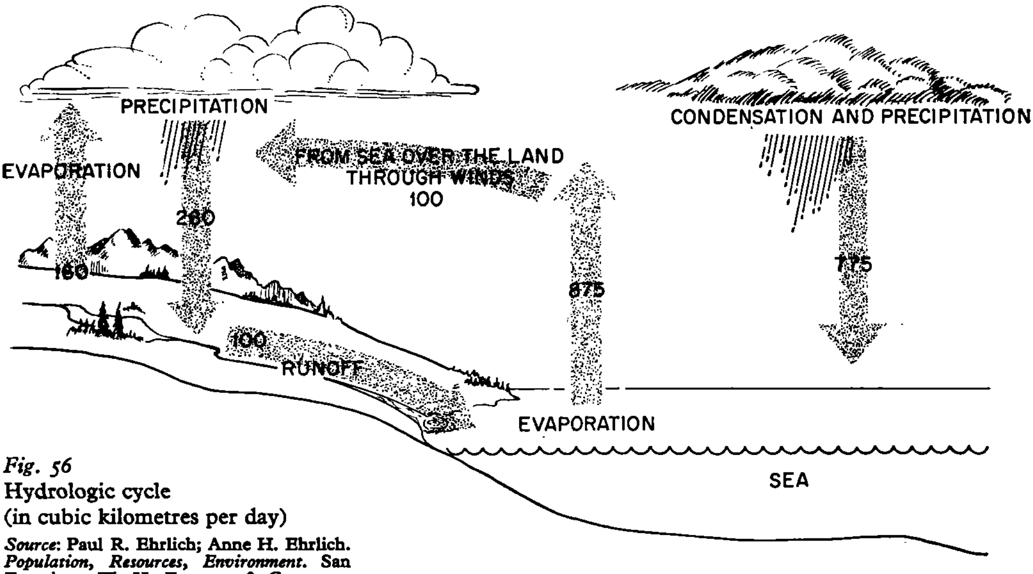


Fig. 56
Hydrologic cycle
(in cubic kilometres per day)
Source: Paul R. Ehrlich; Anne H. Ehrlich. *Population, Resources, Environment*. San Francisco, W. H. Freeman & Company, 1972, p. 76.

Evaporation and transpiration

The movement of water from the surface of the earth to the atmosphere occurs by evaporation. When this evaporation occurs from plants, it is termed evapo-transpiration or transpiration. As Fig. 56 shows, the major source of evaporation is the oceans with a magnitude of 875 km^3 per day, but the process also occurs on land at an approximated rate of 160 km^3 per day.

Several factors influence the rate and intensity of evaporation. Because considerable energy is consumed in the process of evaporation, the availability of solar energy becomes a major contributing factor. Generally, the solar energy reaching the earth's surface is greatest at the equator and decreases with distance from the equator, but the nature of the surface involved may greatly modify this over-all pattern. Other factors are the relative humidity of the overlying air (which may in turn be influenced by the ability of the air to move either vertically or horizontally), and the availability of water at the evaporating surface. When soil is very wet, evaporation may be more than from the oceans, because of a broken surface and a relatively greater surface area, but as the soil surface dries, the rate of evaporation decreases sharply. However, transpiration may continue to be very rapid during daylight hours as long as the root zone is wet. Since transpiration generally occurs through the stomatal openings and these close in decreasing light intensities, evaporation through transpiration occurs generally during periods of high light intensity.

By expressing the moisture in terms of depth over the entire surface area of land and oceans, it is possible to compare the rate of evaporation over land and oceans. Over the entire earth, this figure is about 100 cm per year for evaporation as well as for

precipitation. Over oceans, the average annual evaporation is between 116 and 124 cm, and on land it is around 47 cm. However, since only about one-half of the land surface contributes substantially to evaporation, the average in these areas also approximates 100 cm per year over-all. Evaporation near the equator is generally much higher than at northern latitudes, because of the increased air temperatures affecting the saturation of the atmosphere. In Finland, at 65° N, the average evaporation is 20 cm per year, in south-eastern England (50° N) it is 50, in North Carolina, in the U.S.A. (35° N), it is 80 to 120. In equatorial Africa, e.g. the Congo Basin, evaporation is 120 cm/yr, in Kenya it is 150 cm/yr and in the papyrus swamps of the Nile, in southern Sudan (10° N), it is 240 cm per year. Here the Nile carries its water into the desert where solar radiation is high and evaporation is augmented by hot air that is also very dry.

Condensation and precipitation

The first phase in the return of the water from the atmosphere to the surface of the earth by precipitation is its conversion from water vapour to liquid by condensation. This occurs when the relative humidity of the air reaches 100 per cent, or the dew-point, the water falling or rain, snow or heavy mist (fog).

When air comes in contact with cold objects or surfaces, heat is lost to the object by conduction with resultant formation of dew. A slow mixing of air which has been cooled by conduction may result in fog. If there is very little air movement perhaps only dew will form, while extensive air movement results in rapid dispersal of air such that condensation does not occur. Cooling of the air may also occur in an uplift of moist air. In this case, the air expands with a loss of heat energy and the air temperature is lowered to the dew-point with resultant cloud formation.

There are four major causes of upward movement of air and cloud formation. Air may be heated by conduction from the earth's surface with resultant convection currents in the atmosphere producing a towering column of water droplets and turbulent air called a cumulus cloud. The second cause of upward air movement is the movement of air over rising ground, notably in mountain ranges. The result is often a wet climate on the windward side of mountain ranges. As the air passes over the mountain range, it descends with decreasing ground elevation. This results in cloud dissipation and in the case of a dry climate of rain 'shadows' characteristic of the leeward side of the mountain range. The third cause is that of a convergence or collision of two air masses of different temperature or density. The warmer, lighter air rises above the colder, denser air. When the action is initiated by the colder air, the uplift may be quite dramatic with resultant turbulent weather and the cumulus type clouds of a cold front. When initiated by the warmer air, the action is likely to be more gentle with condensation over a wider area with resultant cirrus and then stratus clouds forming. Precipitation in this case is likely to be more widespread and less intense than with a cold front.

A fourth cause of uplift of air masses is the formation of low pressure areas in the upper atmosphere. The resultant uplift results in condensation.

Precipitation results from condensation, but condensation does not necessarily lead to precipitation. Clouds may form and dissipate without precipitation, but when the moisture droplets (or crystals) reach the size where the air can no longer support them, then precipitation occurs.

Distribution of precipitation

The amounts and distribution of precipitation have major influence upon structure and function of natural ecosystems as well as human settlement. Furthermore, although there are many regional and local modifications, the prevailing pattern of circulation of the atmosphere constitutes a major influence upon distribution of precipitation. Fig. 57 shows a highly simplified diagram of the general circulation of the atmosphere. As the trade winds move from cooler latitudes towards the equator, moisture is picked upon and deposited in the equatorial region. This results in some fairly dry regions north and south of the equator, but a very wet equatorial region. Westerly trade winds, on the other hand, move from warm to cool climates. Subtropical anticyclones are a series of high pressure areas which stimulates easterly flow on the equatorial side and westerly flow on the poleward side. The areas are displaced northward and southward due to the orbital movement of the earth.

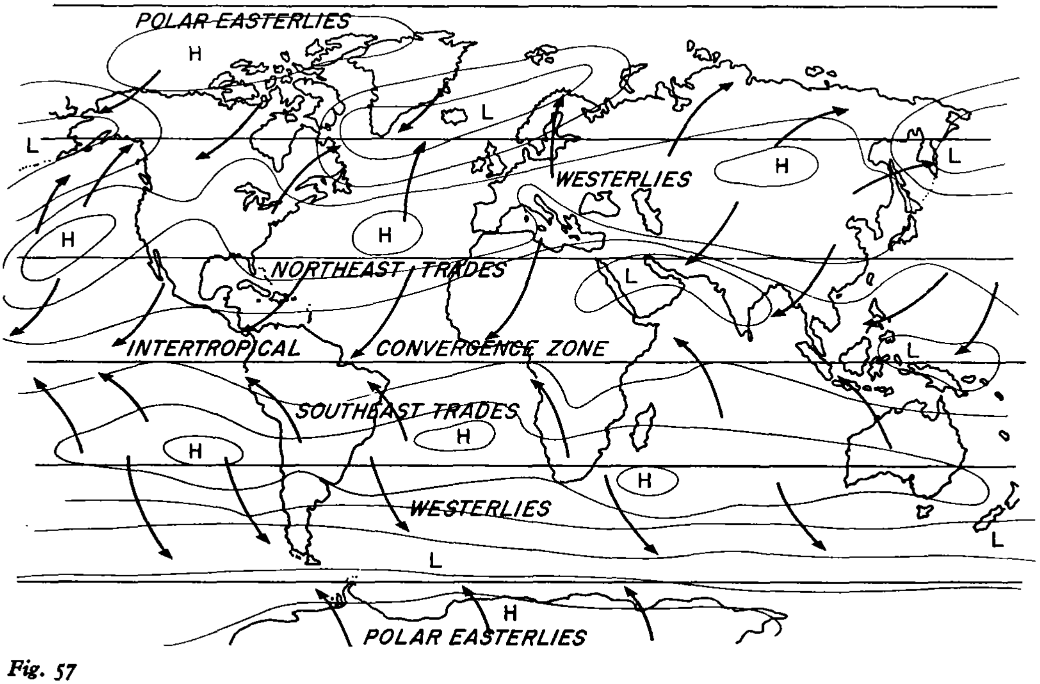


Fig. 57
The general circulation of the atmosphere

Source: Earth Science Curriculum Project. *Investigating the Earth*. Boston, Houghton Mifflin Co., 1967, p. 167.

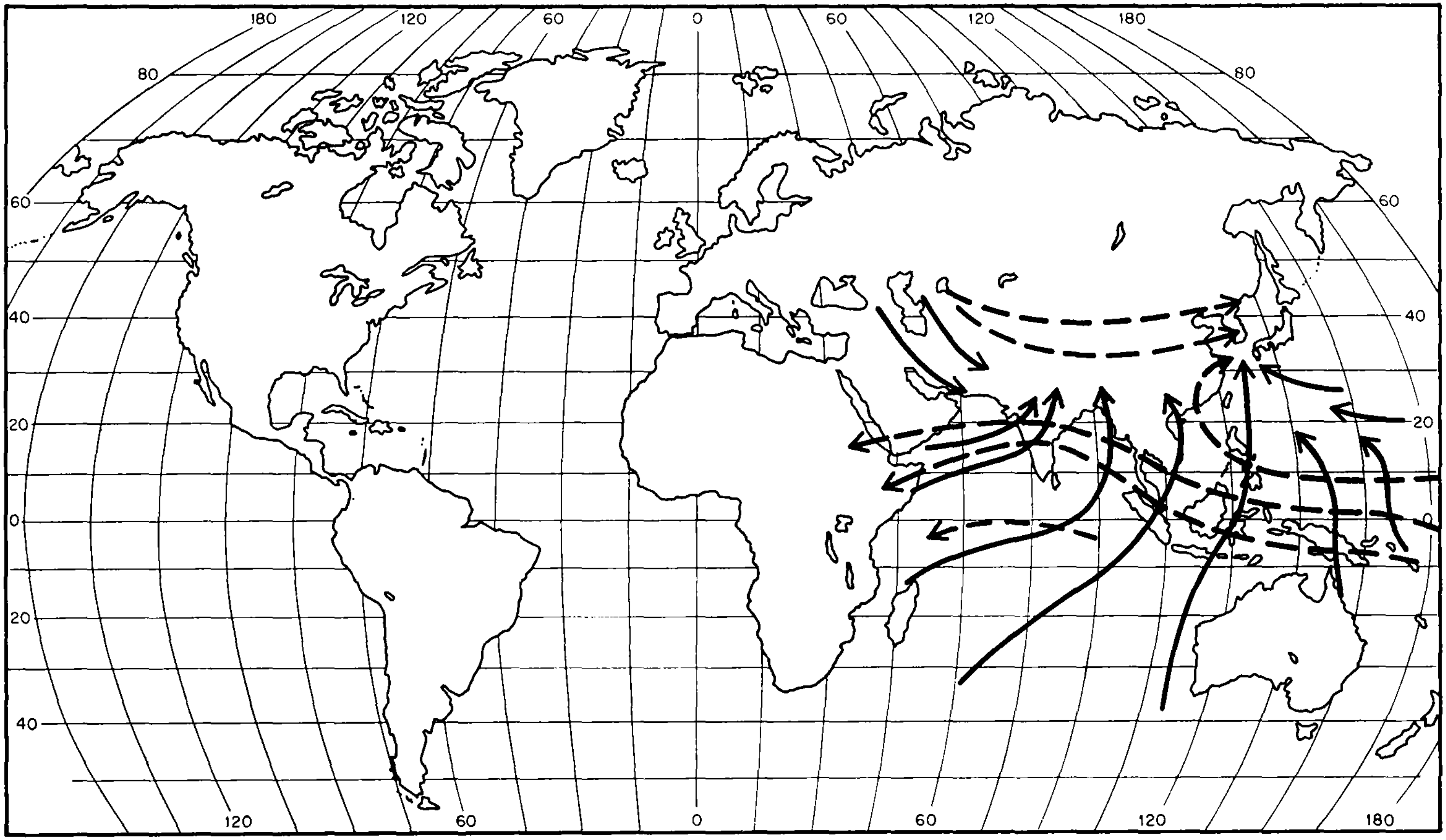


Fig. 58
Summer Asiatic monsoon circulation

Source: Franklin W. Cole. *Introduction to Meteorology*. New York, Wiley, 1970, p. 228.

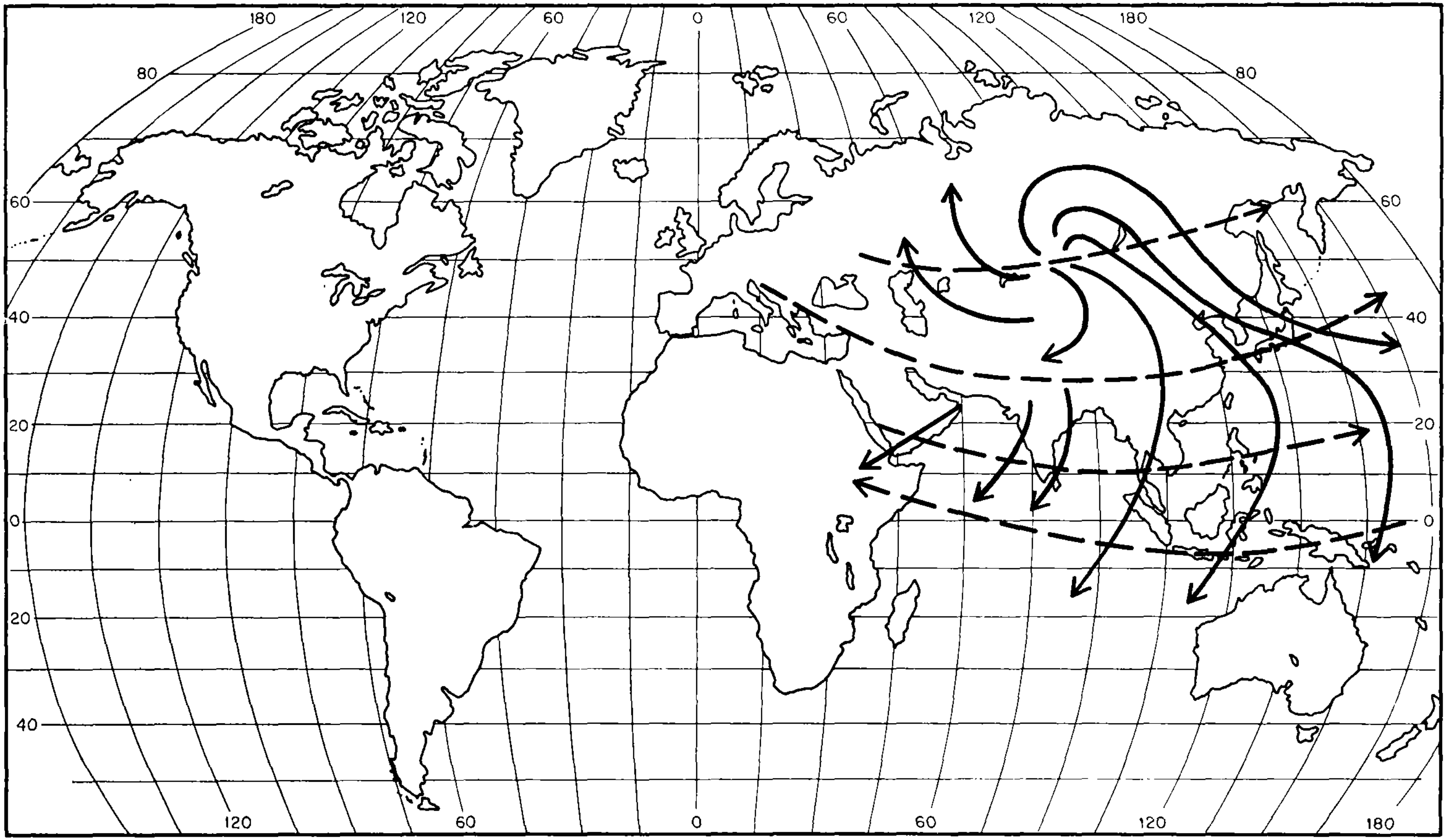


Fig. 59
Winter Asiatic monsoon circulation

Source: Franklin W. Cole. *Introduction to Meteorology*. New York, Wiley, 1970, p. 229.

Secondary circulations modify the general circulation of the atmosphere as shown in Fig. 57. These secondary circulations are caused by variations in energy balance which are related to energy reflection and absorption at different points on the earth's surface. In the hot seasons land temperature rises more rapidly than water temperature. The result is that air over land is warmed more than air over water and thermal low pressure areas are established over land. In the cold seasons the temperature change is again more rapid over the land and the establishment of dense, cold air over land tends to result in a thermal high pressure area. These are the seasonal, monsoon circulations. For example, during the winter, the eastern part of the Asian continent is characteristically covered by a high pressure area centred in Siberia, though this is replaced during the summer by a low pressure area. During the winter, the cold, dry air (called the northeast monsoon) pours over all the South-East Asia from the continent providing fair weather for the region. During the summer, the hot, moist air from the Indian Ocean moves north-east across India, South-East Asia, and the Philippines, bringing the cloudiness and rain of the south-west monsoon. Fig. 58 shows the summer Asiatic monsoon circulation, and Fig. 59 shows winter Asiatic monsoon circulation. Monsoons have also been described for Northern Australia, Africa, Chile, Spain and the United States.

Further modifications of these broad patterns of air circulation occur at regional and local levels in land and sea breeze circulation, valley and mountain breezes and air movements involved in cloud formations already discussed. The point here is that the world-wide general precipitation pattern is a result of numerous interacting factors; among these, land and sea energy relationships, air current and topography are particularly significant.

In relation to the above considerations, it is instructive to consider the pattern of annual mean precipitation on a world-wide basis. Fig. 60 provides information on this pattern. Here, the high precipitation band along the equator, referred to earlier in connexion with the convergence of the trade winds, is evident, as well as the comparatively dry regions 20–30° north and south of the equator. See Fig. 61 for average rainfall during the south-west monsoon season.

Water and living things

A major proportion of the fresh weight of living organisms is water. It functions as a medium or substrate in physiological processes. Then hydrogen of water forms part of food molecules during photosynthesis, and these molecules in turn are modified and re-synthesized into various structural components which serve as foods for other organisms. However, most of the water entering green plants is not used in photosynthesis, but passes through the plant in transpiration. In both plants and animals, water is also the major medium of internal transport of certain substances essential for life.

Various species have become adapted to various climates, including those of limited rainfall, but the availability of water is often a limiting factor in the distribution

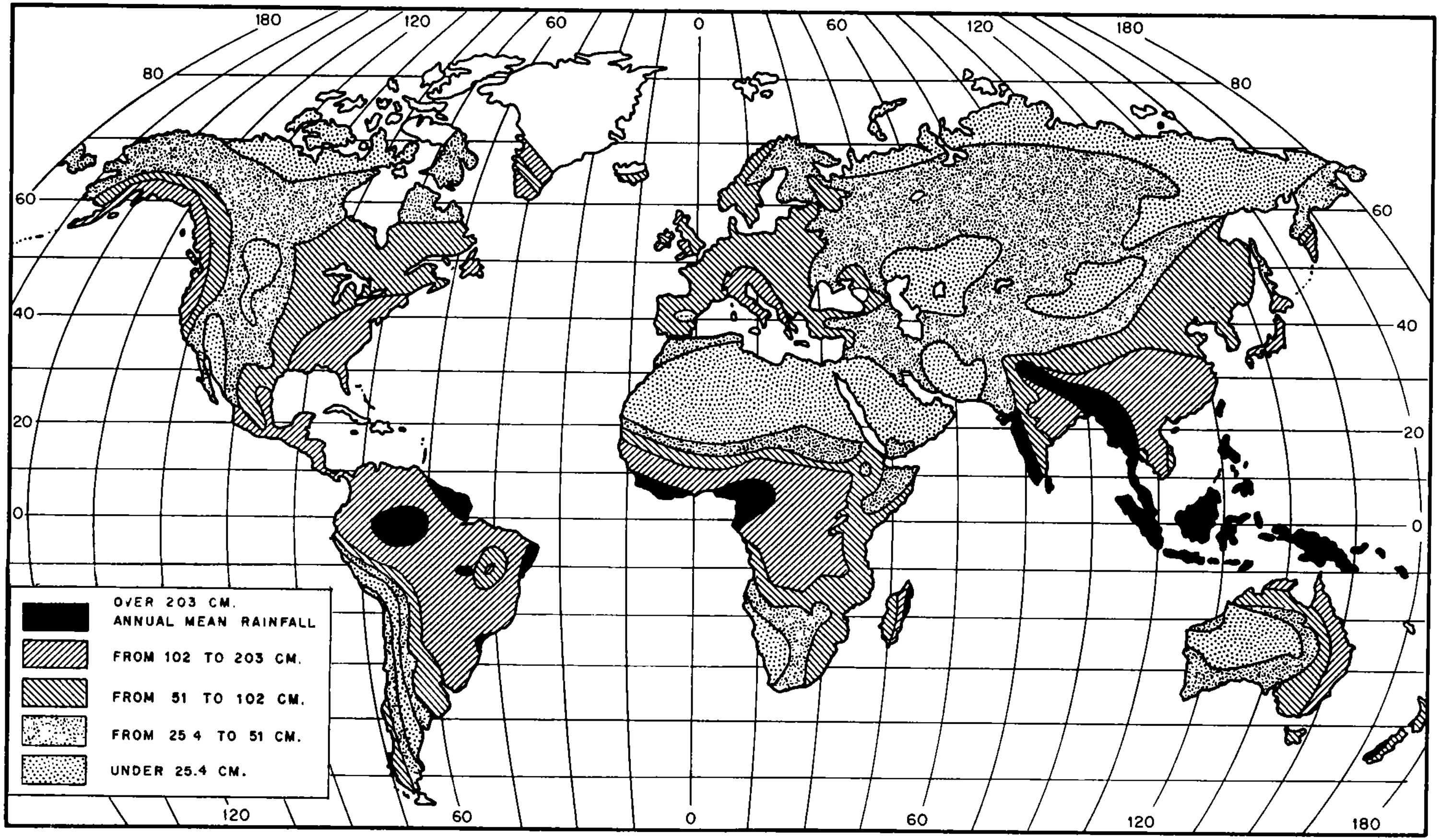


Fig. 60
The pattern
of world precipitation

Source: Edward J. Kormondy. *Concepts of Ecology*. Englewood Cliffs, New Jersey, Prentice-Hall, 1969, p. 38.

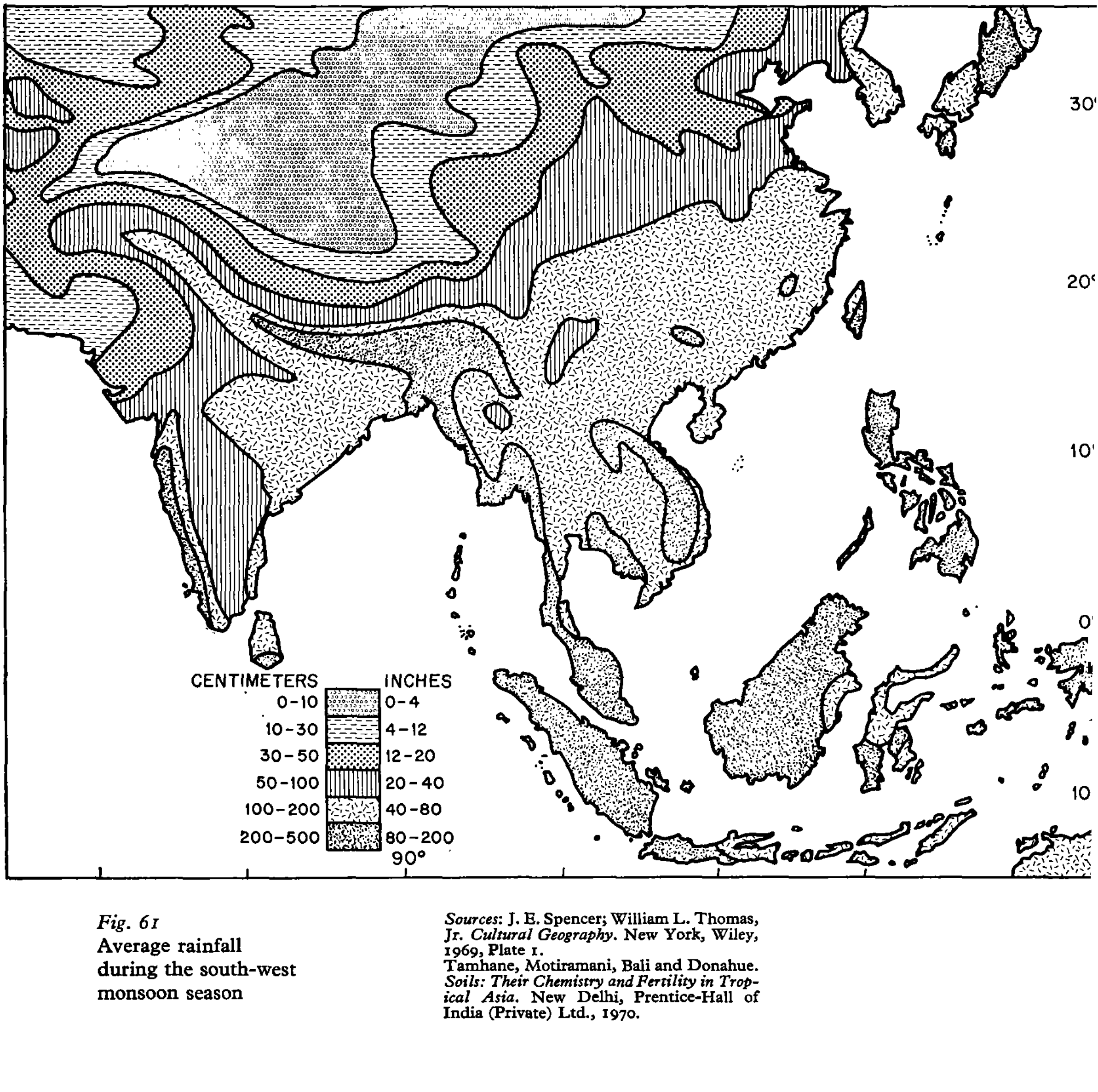


Fig. 61
Average rainfall during the south-west monsoon season

Sources: J. E. Spencer; William L. Thomas, Jr. *Cultural Geography*. New York, Wiley, 1969, Plate 1.
Tamhane, Motiramani, Bali and Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia*. New Delhi, Prentice-Hall of India (Private) Ltd., 1970.

of a species and, indirectly, in the pathway of its evolution. Examples of such species include the cacti, succulent plants, mangrove, the jerboa, the camel and millipedes.

It is interesting to note here the cultural adaptation of the people of Rajasthan to limited water supply. Climatically the desert of Rajasthan is extremely arid with an annual rainfall not exceeding 10–13 cm. Any rain that falls percolates through the sandy desert soils to the water table. Thus in order to irrigate their lands, the people of Rajasthan have resorted to windmills to bring water up to the surface.

From the above, it is obvious that population density, community and ecosystem structure are influenced by the availability of water. Terrestrial ecosystems are generally

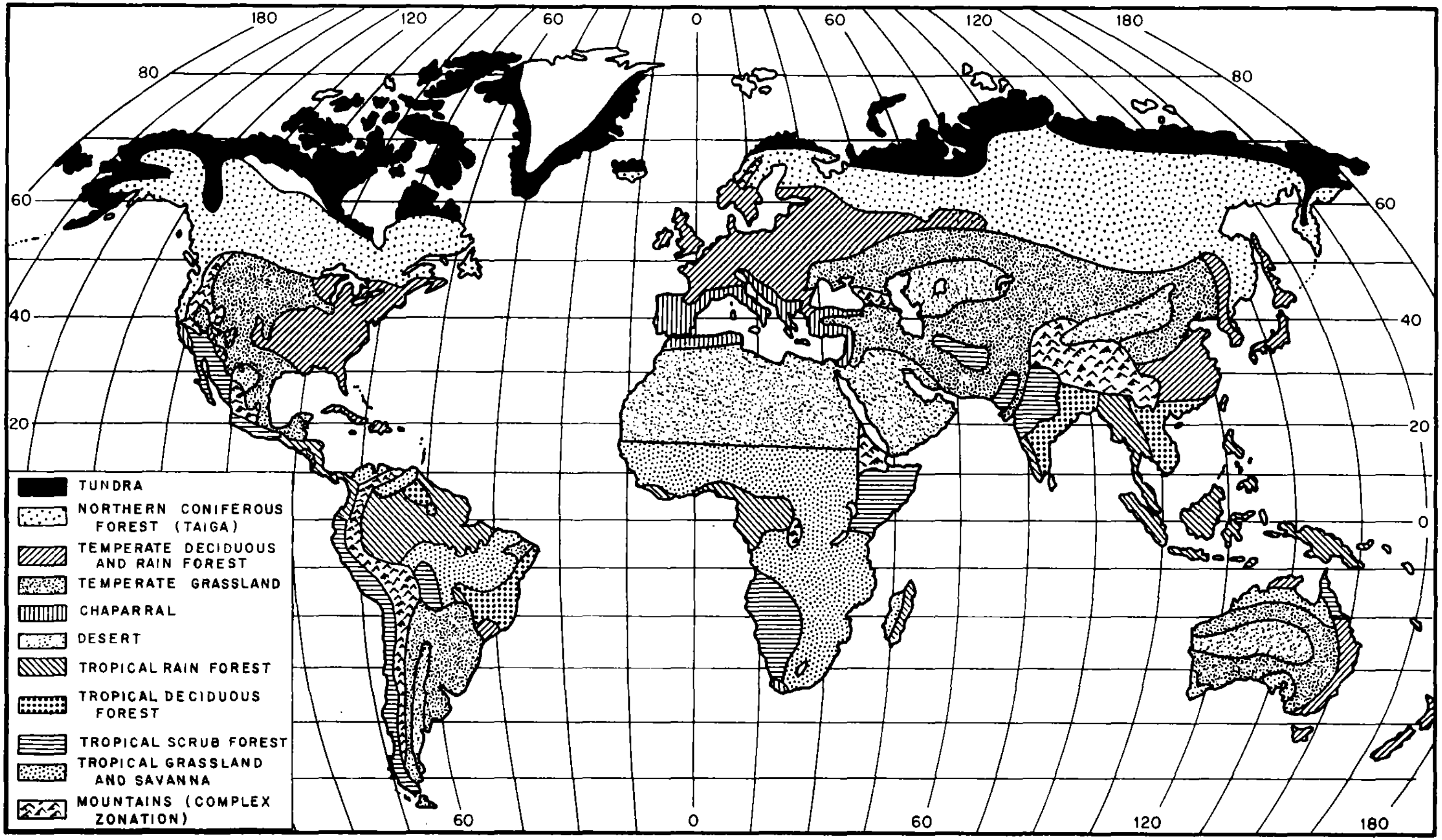


Fig. 62
Major biome types of the world
Source: Edward J. Kormondy. *Concepts of Ecology*. Englewood Cliffs, New Jersey, Prentice-Hall, 1969, p. 114.

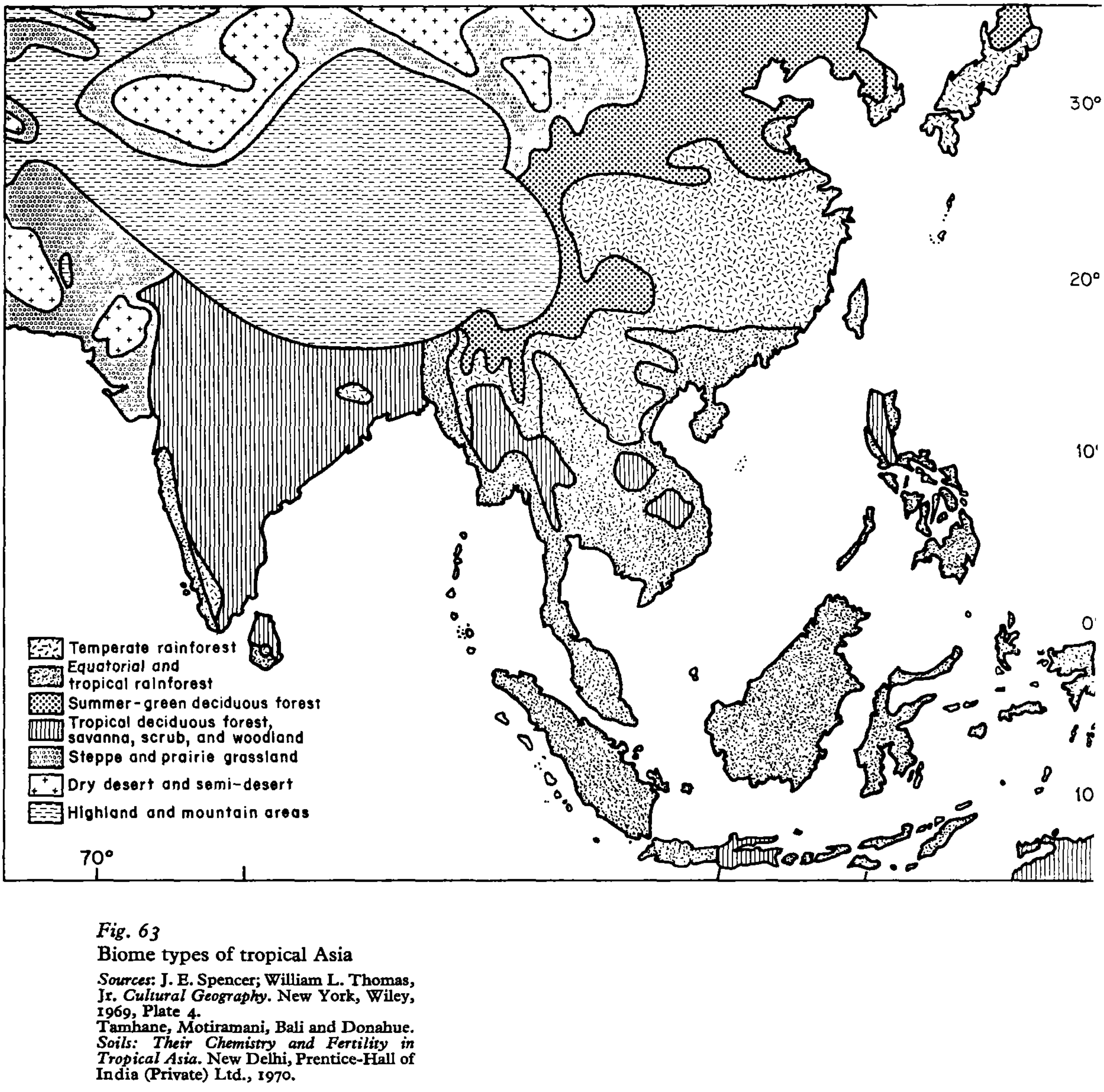


Fig. 63
Biome types of tropical Asia
Sources: J. E. Spencer; William L. Thomas, Jr. *Cultural Geography*. New York, Wiley, 1969, Plate 4.
Tamhane, Motiramani, Bali and Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia*. New Delhi, Prentice-Hall of India (Private) Ltd., 1970.

classified on the basis of the predominant types of vegetation and are called biomes. Fig. 62 shows the distribution of the major biome types. It is clear that biome similarities and differences have considerable latitudinal orientation. When it is recalled that solar radiation (and, therefore, generally temperature) is directly associated with latitude and that wind patterns (and, therefore, generally precipitation) are also associated with latitude, the key roles of energy and water to biome types and productivity become evident. (See Fig. 63 for biome types of tropical Asia.)

Water quality

Pure, uncontaminated water does not occur in nature and the quality varies widely with natural conditions and human activities. When compared to man's interaction with the atmosphere, the water environment presents a greater challenge for better management because of the uneven distribution of usable water over the earth. Indeed, water management problems can, in some parts of the world, far exceed the application of any form of modern technology, yet we have been familiar with the physical, chemical and biological processes associated with natural waters for many years. But, in spite of the vast amount of knowledge, we are far from a complete understanding of the phenomenon of water, hence many techniques that are used today in preparing water for domestic and industrial consumption are highly empirical.

Natural impurities

Natural impurities in water include both chemical and biological materials which may be either in solution or dispersed. Impurities are not essentially dangerous or even undesirable for human consumption. Nevertheless, water is considered contaminated by one or more substances, if it is not suitable for any of the following uses; domestic, agricultural or industrial water needs, fisheries operations, particularly in freshwaters, wildlife conservation and recreational uses, as in swimming pools, boating lakes, etc. Many substances constitute chemical impurities since they contribute molecules that are not water molecules. Oxygen, nitrogen and carbon dioxide may be picked up even during rainfall and are generally dissolved in water to some extent though not themselves dangerous to living systems. Soluble salts of calcium, magnesium and sodium, which are natural constituents of water following its contact with soil or rock are not harmful in their usual concentrations but may cause other difficulties. The bicarbonates, sulphates and chlorides of the above three metals make up almost all of the mineral salt impurities in natural water supplies and account for a condition known as 'water hardness'. When common soaps (but generally not synthetic detergents) are used with certain waters, a curd or precipitate forms. The salts are also deposited in water pipes,

Table 22. Sample natural water analysis

Temperature	17.8° C	Alkalinity	71 milligrammes/litre
Dissolved oxygen	8.2 milligrammes/litre	Hardness	121 milligrammes/litre
pH (hydrogen ion activity)	8.0	Colour	7 (units)
Biochemical oxygen demand	5.2 milligrammes/litre	Turbidity	35 (units)
Chemical oxygen demand	19 milligrammes/litre	Sulphate	70 milligrammes/litre
Chloride	20 milligrammes/litre	Phosphate	0.1 milligramme/litre
		Total dissolved solids	200 milligrammes/litre
		Coliform bacteria	100 per 100 millilitres

Source: 'Water Pollution Surveillance System. Annual Compilation of Data. Oct. 1.1962–30 Sept. 1963'. Public Health Service Publication No. 663, Vol. 2. Washington, D.C., Government Printing Office, 1963.

water heaters, steam boilers and cooking utensils, causing inconvenience and maintenance problems. Table 22 presents a sample natural water analysis, the samples being taken at random, indicating the form and magnitude of water quality factors that are generally determined. Other data would include trace metals, organic compounds, algal and plankton populations.

More serious problems arise when the water containing salt contaminants is used for the irrigation of soils. Salt concentrations allowable in irrigation vary with such factors as the crop under cultivation, quantity of water available, effectiveness of drainage, temperatures and evaporation rates. Allowable salt concentrations in irrigation water range from 70 to 3,500 milligrammes per litre, but water in the higher portion of this range should not be used in regions where evaporation rates are high and drainage is poor. As the water, often already available in limited quantities, evaporates, the salts remain at or near the soil surface in increasing concentrations. The result is an increased osmotic potential of the soil solution with decreasing availability of the water to plants and the resultant retarded plant growth. Large areas of productive land in various parts of the world (including Africa, India, Pakistan and Australia) have been rendered unproductive through unrestricted irrigation using water too high in natural salt concentrations.

Human contaminants and sources

Salt contamination. As human populations increase and as commercial and industrial enterprises develop, numerous additional sources of water contamination arise. For example, brines (salty water) may be released from oil well borings into normally fresh water. When oil is brought to the surface from a deep well, considerable brine is often brought up with it. Disposal of this brine is a problem and one technique has been to pump the salt water down dry oil wells or down wells drilled for this purpose. Numerous industries such as plastic, steel, wood pulp and chemicals operations have used this technique for the disposing of a wide range of waste materials such as phosphates, nitrates, chlorides, cyanides, phenols, sulphates, chlorinated hydrocarbons, alcohols, acetates, ketones and others. Under certain conditions, this may be a fairly effective method of waste disposal, but numerous precautions to safeguard environmental quality are necessary. Injection must be into a permeable sedimentary rock layer such as sandstone or limestone which is porous enough to hold liquids and transmit them. If these wastes are to be properly contained, the permeable layer must be contained both above and below by impermeable layers and it must occur below the water table for the region. Ignoring these precautions can result in contamination of water supplies used for human consumption or soil irrigation.

Acidity. In areas involving mining exploitation (e.g. coal and copper) acid seepage into water supplies often becomes a problem. As bacteria convert ferrous to ferric oxides, sulphuric acid is released. Effects on water supplies include the addition of iron and sulphate, lowering of the pH and reduction of available oxygen in the water. Such discharge occurs not only from active mines, but also from abandoned mines.

Organic wastes. These materials which enter water supplies constitute a major contributory factor in the human contamination of water. Late in the nineteenth century, the germ-disease theory made possible the development of water bacteriology. It soon became evident that a major source of water contamination was sewage and that many diseases were the result of this water-borne contamination. Many diseases which are water-borne may also be contracted in other ways, but impure water has been a common medium for transmission of a significant number of them. Table 23 gives examples of pathogens' association with water in various ways and some of their related diseases.

Table 23
Pathogens associated with water and diseases caused by them

Pathogen	Disease
BACTERIA	
Bacillus typhosum (Eberthella typhi)	Typhoid fever
B. paratyphosum	Paratyphoid fever
B. dysenteriae	Bacillary dysentery
Vibrio cholera	Cholera
PROTOZOA	
Entamoeba histolytica	Amoebic dysentery
Giardia lamblia (*G. intestinalis*)	Giardiasis
VIRUSES	
Liver-infecting virus	Infectious hepatitis
Polio virus	Poliomyelitis
Virus (spread by mosquito needing water for part of life cycle)	Yellow fever
PLATHELMINTHS and NEMATHELMINTHS	
Tapeworms	Onchocerciasis
Schistosoma	Bilharziasis
ALGAE	
Euglena and other algae	Gastroenteritis

Sand filtration and disinfection of water with chlorine probably contributed most to water-borne disease control, however, and these processes became acceptable in some parts of the world around the end of the nineteenth century and the beginning of the twentieth.

A second common result of organic wastes in water is that the organic matter provides food for bacteria which in turn increases their biological oxygen demand (BOD). The dissolved oxygen (DO) of the water is thus decreased. The decreasing DO of the

Table 24
Comparison of decomposer end products under differing conditions

Aerobic conditions		Anaerobic conditions	
C	CO_2	C	CH_4
N	NH_3+HNO_3	N	NH_3+amines
S	H_2SO_4	S	H_2S
P	H_3PO_4	P	PH_3 and phosphorous compounds

Source: Richard H. Wagner. *Environment and Man*. New York, W. W. Norton and Co., Inc., 1971, p. 112.

water has several effects. Of the major forms of life in freshwater, fish generally have the highest demand for oxygen and are likely to be the most adversely affected. Under extreme conditions, even the aerobic decomposers may be replaced by anaerobic decomposers with a considerable change in the end products. Table 24 compares end products under differing conditions.

Several of the anaerobic end products have characteristic odours, so the change to anaerobic conditions produces quite unpleasant effects. Direct human and animal wastes are a source of organic materials. When the human population density is low, wastes can be buried in the soil and used as fertilizer or if emptied into rivers and streams, wastes would be broken down by aerobic decomposers and the water effec-

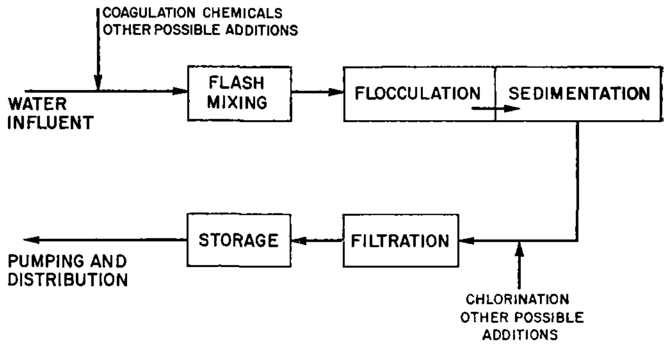


Fig. 64
Major steps in water treatment

tively purified of organic wastes by natural processes. There are several other significant aspects of natural purification. In still water, sedimentation removes suspended matter; some carbon dioxide is removed in waterfalls and rapids; colour is bleached by sunlight and pathogenic bacteria are consumed by other organisms. As human population densities increase and urbanization occurs, these solutions are inadequate because organic wastes from people and farm animals exceed the natural purification capabilities. In rural areas, the use of deep wells alleviates the problem, while high density urban populations require complex water and sewage treatment facilities. (See Fig. 64 showing major stages in sewage and waste water treatment.)

Industrialization and pollution. As society becomes industrialized, again major pollution problems arise, many of which involve organic wastes. Pulp and paper mills produce a waste sulphite liquor containing lignosulphonate and sugars which are often dumped directly into natural and open waters. Increasingly, productive uses for these materials are being found, such as road surface stabilizers, drilling mud additives, vanillin, etc., further minimizing the need for allowing this form of environmental deterioration.

Large amounts of water are used in petroleum refining for cooling and for separating impurities from the desired organic compounds to be refined. Significant quantities of organic materials should not enter water supplies in this way, but poor operating procedures often permit leakages into and consequent contamination of water resources.

Food processing may be a major contributor of organic wastes in some areas. Animal slaughter houses often release water containing large amounts of blood and other waste tissues directly into open waters. Fruit and vegetable processing plants utilize large amounts of water to rinse food material being prepared as well as in the peeling and blanching of the foods. Again, the result is waste water high in organic content.

It must be pointed out that numerous other industrial effluents with harmful effects are routinely released into water supplies. Toxic heavy metals, pesticides, salts and acidic water from mines, and turbidity from steel mills are illustrative examples of additional industry-related water pollution problems.

An additional point related to organic wastes bears emphasis here. As the wide array of organic waste materials is broken by decomposers, carbon dioxide, water, and inorganic nutrients like phosphates and nitrates are released into the environment. These inorganic nutrients are in turn used by autotrophs in the production of more organic food materials, so their presence in nature is not only desirable but essential. Problems emerge, however, when their concentrations are drastically changed. Concentrations of phosphates in water are usually low since most phosphates are quite insoluble in water, with an over-all effect of limiting undesirable algal growth. The ready availability of phosphates removes this natural limiting factor and permits a rapid population growth of algae. At night, the algae require oxygen for respiration and because of the excessive increase of the algal population the oxygen content may become too low to support other populations such as fish. In some areas, the problem of excessive phosphorus is further aggravated by agricultural use of highly soluble fertilizers of which 10 to 25 per cent may be leached away into surface run-off before it is used by plants. Furthermore, in areas where phosphate-containing synthetic detergents are used in household cleaning in general, very large quantities of phosphorus are discharged into natural waters. There is a strong case, in this respect, for encouraging the manufacture of citrate-based detergents, which do not have such adverse breakdown effects.

Similarly, increased nitrate concentrations may occur in water supplies from agricultural fertilizer run-off of untreated wastes from high density human or farm animal populations such as from chicken hatcheries or cattle herds. There is evidence that an excessive nitrate content of water is the cause of disease (methaemoglobinaemia) in infants, and, at lower concentrations, chronic developmental deficiencies in infants. It is now apparent that the development of high density human and animal populations inevitably must be accompanied by extensive sewage treatment, if major contamination of water supplies and resulting human suffering are to be avoided.

While some kinds and levels of impurities are not necessarily harmful, water standards with regard to chemical content must be controlled if human health is to be protected. Table 25 shows major contaminants and concentration limits which are extensively utilized. Section 1 includes substances which should not be present if other water sources can be found. Section 2 of the table includes substances and concentrations which should indicate grounds for rejection of the water supply for domestic use.

Table 25
Water and contaminants from industries

1. Substance	Concentration (mg/litre)
Alkyl benzene (ABS)	0.5
Arsenic (As)	0.01
Chloride (Cl)	250.0
Copper (Cu)	1.0
Carbon chloroform extract (CCE)	0.2
Cyanide (CN)	0.01
Iron (Fe)	0.3
Manganese (Mn)	0.05
Nitrite (NO_2)	45.0
Phenols	0.001
Sulphate (SO_4)	250.0
Total dissolved solids	500.0
Zinc (Zn)	5.0

2. Substance	Concentration (mg/litre)
Arsenic (As)	0.05
Barium (Ba)	1.0
Cadmium (Cd)	0.01
Chromium (Cr^{6+})	0.05
Cyanide (CN)	0.2
Lead (Pb)	0.05
Selenium (Se)	0.01
Silver (Ag)	0.05

Source: P. Walton Purdom (ed.). *Environmental Health*. New York, Academic Press, 1971, p. 154.

Earlier in this chapter, it has been noted that pure water as such does not occur in nature; nevertheless there are many places over the earth's land mass where natural waters of high quality may be found and these continue to be used by human communities, providing the main or only source of water needed for sustenance and survival. Spa waters and certain spring waters have been shown to possess properties of useful, medicinal value and these are not only taken by thousands of persons who may visit such sources, but they are also bottled and distributed commercially to many other centres for human consumption. Yet the peculiar properties of natural waters, in so far as soluble ingredients are concerned, are not always with desirable qualities for the support of living systems in general and human populations in particular.

Examples of the adverse effects of increased concentrations of inorganic chemical substances in water supplies that may be drawn on for various human activities are given in Table 26.

Pesticide residues and water. Pesticides are man-made substances. Their introduction into the environment, therefore, is one of concern, particularly because of the persistent effects of pesticide residues in soil, vegetation and water; among these increasingly numerous products, the chlorinated hydro-carbons, such as DDT, have been studied

Table 26. Water-soluble inorganic chemical substances entering municipal water supplies from natural sources and their adverse effects upon domestic and industrial operations

Substance	Natural source	Adverse effect
Silica Si	Clay minerals, opal, rock minerals	Forms scale on boilers and steam turbines, inhibits pipe corrosion
Iron Fe	Igneous and sandstone rocks, iron pipes, pumps, storage tanks, etc.	Stains plumbing fixtures, laundry and cooking utensils, spoils water taste and colour
Manganese Mn	Soils and sediments, metamorphic and sedimentary rocks	Has undesirable taste; leaves deposits on food during cooking; stains plumbing fixtures and laundry
Calcium Ca^{++}	Gypsum, calcite, clay, limestone, rock minerals	Combines with other minerals to form scale in boilers; inhibits the formation of soap suds
Magnesium Mg^{++}	Limestones, clay, rock minerals	Same effects as calcium
Sodium Na^{+}	Clay, sediments, industrial wastes, rock minerals	Produces scale and corrosion in boilers; combines with potassium carbonate to cause wood deterioration
Potassium K^{+}	Micas, clay, rock minerals	Same effects as sodium
Bicarbonate $(CO_3H)^-$	Limestones	Combines with other minerals to form scale in pipes
Carbonate $CO_3^=$	Limestones	Same effects as bicarbonate
Sulphate $(SO_4)^=$	Oxidation of sulphide ores, sulphate minerals, industrial wastes	Forms scale; causes bitter taste: may be cathartic
Chloride Cl^-	Sedimentary and igneous rocks, salty water forced upstream into tidal estuaries	Has a salty taste, can be harmful to health
Fluoride F^-	Rock minerals, fluorite, mica	Increases resistance to tooth decay but in excess may cause mottling of tooth enamel
Nitrate NO_3^-	Atmosphere, legumes plant debris, animal excrement, nitrogenous fertilizers, sewage	Has a bitter taste, harmful in excess, especially to infants
Calcium carbonate[1] $CaCO_3$	Limestone (a standard measurement for the hardness of water)	Inhibits formation of soap suds, forms an unsoluble scum or curd in washing machines

1. See carbonate and bicarbonate.

Source: L. B. Leopold; K. S. Davis and the editors of *Life Water*. New York, Life Science Library, Time, 1966, p. 193.

fairly comprehensively and the data available on them allows for the establishment of recommended safeguards and control measures, as far as environmental contamination is concerned. When water is affected by the accumulation of pesticide residues from soil, vegetation and run-off, then we are presented with a situation in which the health and well-being of fish, birds and other wildlife, associated with water as well as the issue of human health, requires careful analysis and control.

Fortunately, excellent and reliable techniques for the identification and assessment of pesticides and their residues are available and in use. The current need is for public

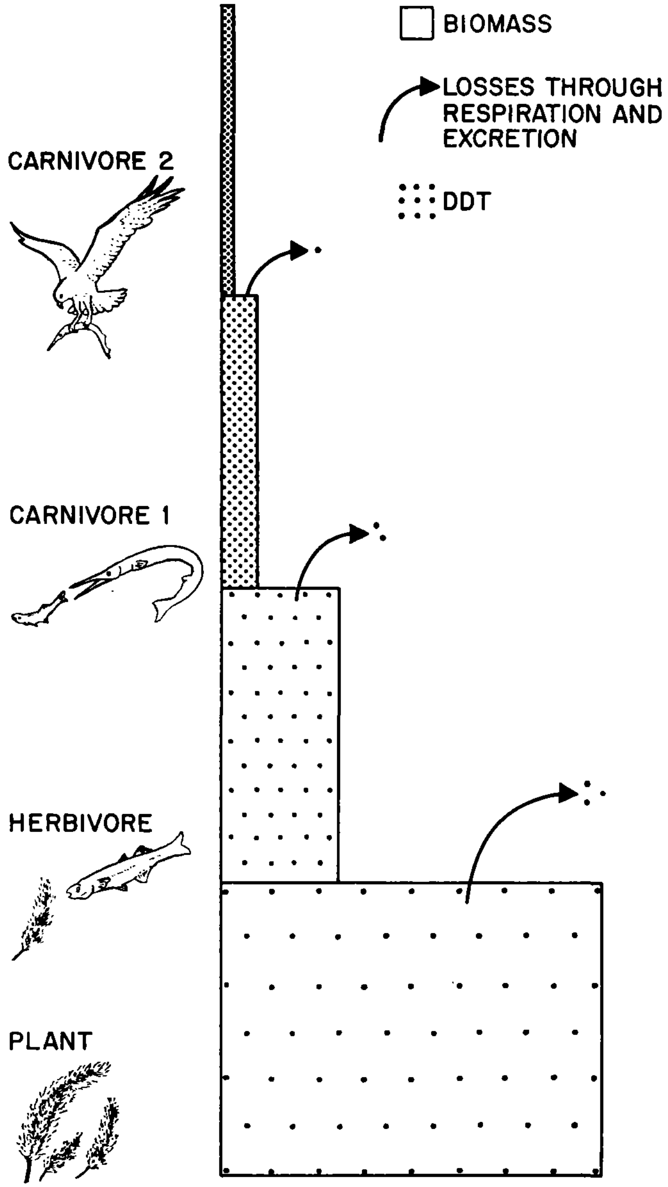


Fig. 65
DDT in a food chain
Source: George M. Woodwell. Toxic Substances and Ecological Cycles, *Scientific American* (New York), Vol. 216, No. 3, March 1967, p. 30.

Note. Concentration of DDT residues being passed along a simple food chain is indicated schematically in this diagram. As 'biomass', or living material, is transferred from one link to another along such a chain, usually more than half of it is consumed in respiration or is excreted (arrows); the remainder forms new biomass. The losses of DDT residues along the chain, on the other hand, are small in proportion to the amount that is transferred from one link to the next. For this reason high concentrations occur in the carnivores.

awareness of the hazards arising from the careless use of such substances and of the consequent adverse effects upon water quality and the good health of man and other living systems.

The chief hazard elicited by such substances in water is that of the concentrating effect in a food chain. DDT in lake water, for example, may be taken up selectively by plankton, which would be eaten by small fish and these would be eaten by larger fish, say *Tilapia*. The concentration of the residue is enhanced at each stage throughout the food chain and, although this may not have an adverse effect on humans, it does so in fish and wildlife (see Fig. 65 illustrating the concentration effect of DDT in a food chain involving the water environment).

Radiation and water. For human populations, in general, significant amounts of radiation exposure are related to natural background sources and the widespread use of X-radiography in the hospitals and clinics. Recalling the comments in the introduction to Chapter 5 on 'atmosphere', background radiation also emanates from cosmic rays and the radioactive substances existing in soil, water and the atmosphere. The average background level is between 100 and 125 millirems per annum (1 millirem = 1/1,000th of a rem, or roentgen equivalent man, reflecting the amount of radiation absorbed by human tissues as well as the quality and type of such radiation). Whatever the source, be it natural or man-made, the natural waters of the earth are either directly or indirectly affected. Modern water analysis systems should include a monitoring and control of radiation levels as a further assurance of quality standards. Developing countries must pay as much attention to this issue as industrialized and technologically advanced nations.

Water purification

Clean water for human consumption must be provided if human health is to be protected. In rural low-density situations, underground water supply may be adequate. If the water table is fairly deep, such supplies are usually free of surface contaminants (since the soil functions as an effective filter) and the major task is to prevent contaminants from entering wells for example through their openings at the surface. Bore-holes and covered wells with pipe fixtures are thus preferable to open hand-dug wells. In addition, seepage of human and animal excrement and refuse into the well near the surface is a serious source of contamination and a health hazard. Unless population density is very low, surface water is generally unacceptable for human consumption since natural processes of purification at the surface are easily overloaded. Hence, as human population density increases and urbanization occurs, water purification is necessary to protect human health. The effective provision of pure water generally involves two major aspects: water treatment and sewage treatment. The first of these actually prepares the water for human consumption and the second reduces the impact of sewage on water supplies. Inadequate domestic and industrial sewage treatment thus makes adequate water treatment more difficult.

Sewage treatment

A sewage system should include covered sewage drains from human dwellings and road transported sewage to a central sewage treatment plant. Storm sewers for rain-water run-off should be kept separate from domestic sewers since rain-water causes great variation in volume which may exceed treatment plant capacity. The major aspects of sewage treatment are usually grouped according to primary, secondary and tertiary treatment. Additional aspects which require attention are disinfection and sludge treatment. Fig. 66 provides a summary of treatment processes.

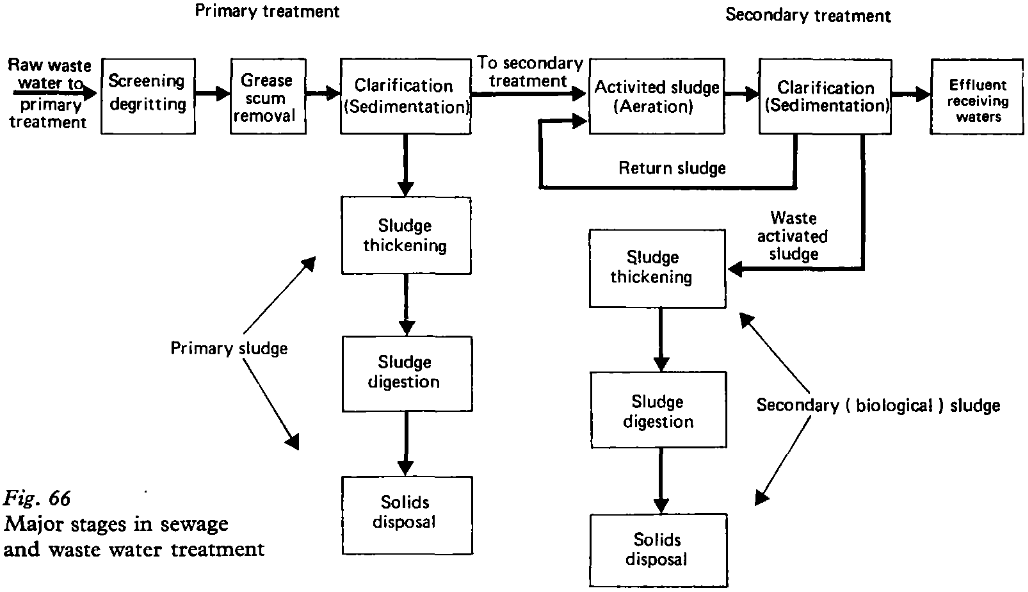


Fig. 66
Major stages in sewage and waste water treatment

Note. There are variations practised on the above scheme, but the process normally begins with the removal of large and fine solids, oily substances and scum, followed by sedimentation and the organic matter is then broken down by micro-organismal activity, with oxidation, eventually with a discharge to open waters of the clarified effluent.

Preliminary and primary treatment. Preparation for the major aspects of treatment is accomplished through preliminary treatment. The three most common components of preliminary treatment are racks and bar screens for removal of large organic solids, grinders or cutters to grind large particles for easier handling in later stages of treatment, grit chambers for removal of sand and grit to decrease potential damage to equipment in later treatment processes.

Primary treatment is the most common and basic form of treatment used. The major features of primary treatment and their functions are: large tanks for sedimentation of solids, scrapers at the bottom of the tanks for removal of settled solids to a sump, pumps for periodic removal of settled solids (sludge) from the sump,

surface scrapers and hoppers on the sedimentation tanks for removal of light solids, oil and grease.

In plants using only primary treatment, the sludge is then treated and the water is discharged after disinfection to reduce bacterial content. Approximately 50 per cent of the influent solids are typically removed using primary treatment.

Intermediate and secondary treatment. When more effective waste removal is desired, intermediate (chemical) treatment and/or secondary (biological) treatment can be used. Chemical treatment requires chemical feeders and mixing units for the addition of chemicals to the sewage to facilitate coagulation and precipitation of solids. Organic waste removal efficiencies run as high as 85 per cent using this method. Chemicals suitable as coagulants include calcium hydroxide, aluminium sulphate and ferric chloride.

Biological treatment utilizes bacteria to oxidize organic materials of sewage to inorganic substances. The principal requirements for effectiveness of this process are suitable pH, high oxygen availability and surface area contact between sewage and bacteria. Four techniques by which these requirements can be met are in use at the present time: oxidation ponds, sand filtration, trickling filtration and activated sludge.

Oxidation ponds are shallow, earth-bottomed ponds or series of ponds in which waste is broken down aerobically (at the surface) and anaerobically (below the surface). Waste water must be detained for a period of about thirty days. Where land cost is low and the construction and maintenance is easy the requirement of only relatively unskilled operations makes this an attractive choice for small municipalities.

Sand infiltration is also quite suitable for small communities. The intermittent sand filters provide physical filtration as well as a medium for an aerobic bacterial population. As the sewage passes through, organic materials are oxidized.

A trickling filter is a bed of 5 to 10 cm stone over which the sewage is sprayed. The bacterial colonies develop on the stones, and as the sewage trickles across, aerobic oxidation occurs. For increased efficiency, a series of such filters is sometimes used.

Activated sludge uses the sewage itself as the medium in which the biological activity occurs. Numerous variations of this method occur, but the procedure generally involves an aerated channel through which the organic matter passes. The organic matter is absorbed by sludge particles which include bacteria and aerobic oxidation occurs.

Tertiary treatment. After secondary treatment, the effluent still contains considerable phosphates, nitrates and other inorganic substances. The purpose of tertiary treatment is to remove such substances. Techniques include algal growth in shallow tanks to remove the nutrients, and addition of lime or alum to remove the phosphorus as a precipitate. Other techniques are also used, depending upon specific components of the effluent (especially from industrial wastes) and requirements of the situation.

Sludge treatment. Treatment and disposal of the sludge which remains after primary and secondary treatment are problem areas. Various techniques are used to provide

further decomposition of the solids. Commonly the material is then dried after which it may be sold for fertilizer, used for land fill or dumped.

Sediments and turbidity. When finely divided suspended particles become fairly concentrated, light is less effectively transmitted through water. This often results in inability of submerged aquatic plants (*Elodea, Chara, Potamogeton*) to carry out photosynthesis. Death of these plants and their decomposition in turn results in reduced dissolved oxygen in water and resultant changes in animal populations (since species require differing concentrations of dissolved oxygen). In addition, the suspended particles eventually settle out, creating a coating of silt on the bottoms of lakes and streams. The habitats of bottom-dwelling populations (such as insect larvae), are thus modified and sometimes destroyed, resulting in elimination of the populations and modification of food webs.

The sources of suspended particles are numerous and not all are controllable through sewage treatment. For example, accelerated erosion and resulting increased sediments may be due to poor land management practices, but other sources like industrial and municipal wastes can be effectively controlled through waste treatment systems.

Organic effluents. These materials are the most common of the water polluants, coming from human wastes and industrial sources such as pulp and paper mills, food-processing plants, petroleum refineries and tanneries. The materials are usually not directly toxic to aquatic life, but their indirect effects may be just as hazardous.

The mechanism which results in ecosystem damage is as follows: the organic wastes released into the body of water serve as a food source for populations of bacteria. The availability of large quantities of food materials permits bacterial population growth. As these organic materials are utilized in aerobic respiration, oxygen dissolved in the water is consumed. The result is that organic wastes released into bodies of water cause an increased biological demand for oxygen (or biological oxygen demand or BOD). Thus, a high BOD tends to result in lowered levels of dissolved oxygen (DO) in the water.

The concentration of dissolved oxygen (DO level) is of critical importance in ecosystem structure since populations of aquatic organisms differ in their DO requirements. Populations of fish (especially cold water fish) have comparatively high DO requirements, invertebrates generally have lower requirements and bacteria, still lower. Thus the high BOD by bacteria in organic waste situations readily results in a lowering of DO well below the tolerance limits of many desirable populations of aquatic ecosystems.

Fig. 67 illustrates some effects of untreated sewage on an ecosystem.

Avenues towards solution include a use of at least primary (but hopefully secondary) municipal waste treatment, the use of industrial wastes for the production of useful by-products, industrial waste (sewage) treatment, and more careful industrial operating procedures.

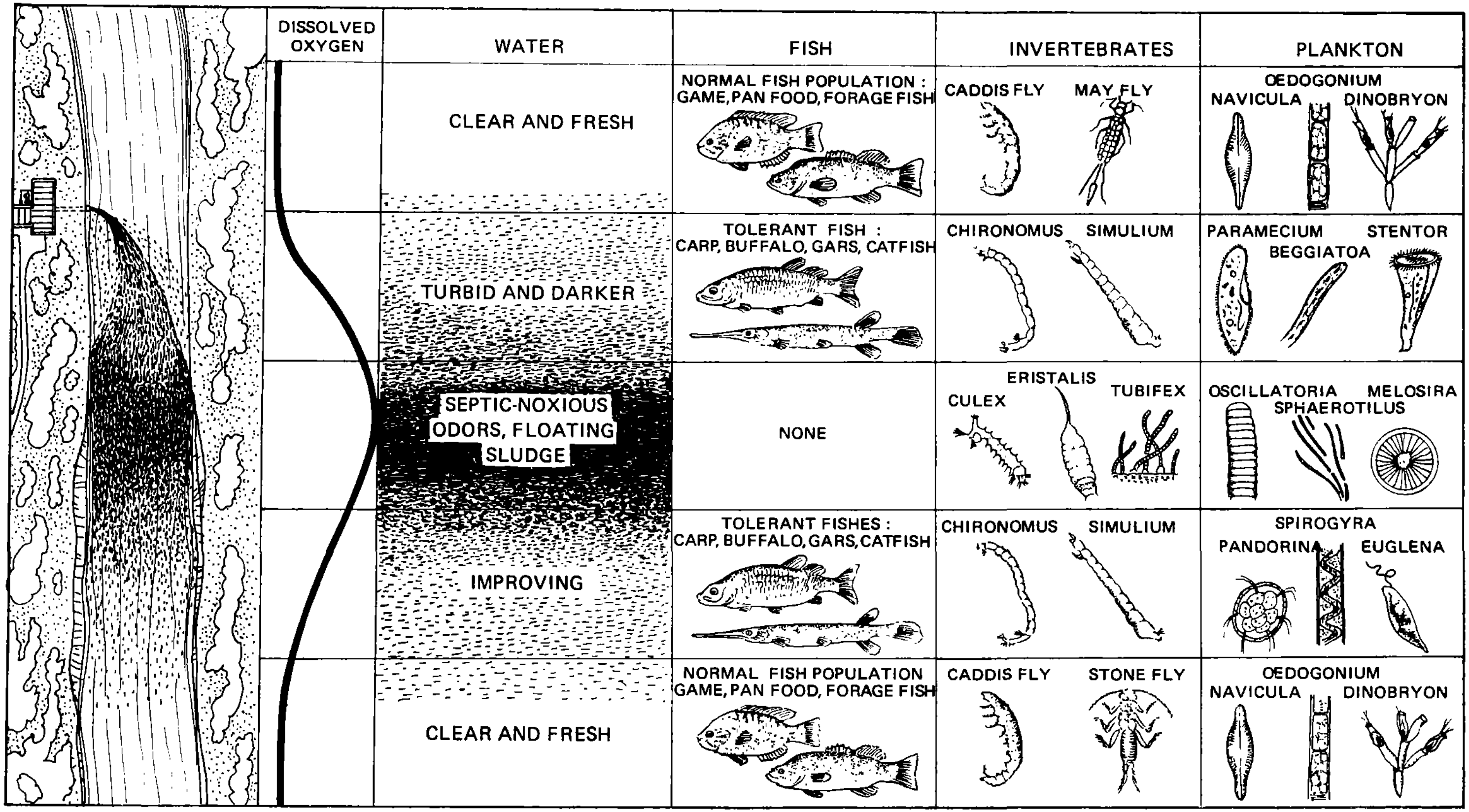


Fig. 67

Some effects of untreated sewage on an ecosystem

Source: Eugene P. Odum. *Fundamentals of Ecology*, 3rd ed. Philadelphia :W. B. Saunders Co. 1971, p. 443.

Note. Pollution of a stream with untreated sewage and the subsequent recovery as reflected in changes in the biotic community. As the oxygen dissolved in the water decreases (curve to the left), fishes disappear and only organisms able to obtain oxygen from the surface (as in *Culex* mosquito larvae) or those which are tolerant to low oxygen concentration are found in zone of maximum organic decomposition. When bacteria have reduced all of the discharged material the stream returns to normal.

Nutrients. Limited availability of nutrients (particularly phosphorus, potassium and nitrogen) is ordinarily a major factor limiting plant growth in nature. When these nutrients become available in abundance, plant growth is greatly enhanced, which is a good reason for adding nutrient fertilizers to agricultural soils.

When organic wastes are decomposed by bacteria, inorganic nutrients remain, and when they are released into rivers and lakes, they function as fertilizers. Especially when phosphate enrichment through these mechanisms occurs, various species of aquatic algae experience a population explosion (or 'bloom'), which further modifies the ecosystem. Sunlight is intercepted and bottom-dwelling plants may be shaded out. Where distinct seasonal patterns exist, these algal populations then die out in late summer resulting in massive decomposition and release of odorous decomposition products (particularly hydrogen sulphide). In some cases, fish kills occur at night during the time of rapid algal growth because of lowered DO due to respiration and lack of photosynthesis at night by the massive populations of algae living in the water. Thus, the nutrient enrichment considered desirable for agricultural crop growth can result in incidental modification of ecosystems when it occurs in rivers, lakes and streams.

Human organic wastes are a major source of nutrients such as the above. Other sources of varying significance are run-off of leaching or agricultural fertilizers, phosphates from household detergents (in some technological societies) and in some countries, animal wastes.

Water treatment. Industrial and municipal sewage treatment usually reduces pollution of water supplies but does not provide water that is safe for human consumption. The purpose of water treatment is to provide safe and palatable water. The major components of a typical water treatment system have been presented in diagrammatic form in Fig. 64, page 154.

Coagulation and aeration. The first step is to add coagulation chemicals to the water and mix them in. The purpose of this step is to bring together the fine colloidal particles into larger clumps that will settle more rapidly. Commonly used coagulents are aluminium sulphate, ferric chloride, ferric sulphate and sodium aluminate. Additional substances are sometimes used as coagulation aids. The pH must be regulated for maximum coagulation in relation to the water and coagulation chemicals used. Activated carbon is sometimes also added at this stage to absorb tastes and odours resulting from plant and animal life (often algae) or industrial wastes.

Often, the water is then aerated, a process in which air and water are brought into maximum contact to improve the quality of the water. Specific purposes include reduction of odours and tastes, reduction of gases such as carbon dioxide, hydrogen sulphide, and methane, and addition of oxygen. Among methods used are perforated pipes, inclined planes and nozzles.

After mixing, the water flows to a coagulating or flocculating basin where it is stirred for a time to build up flocs for maximum settling. To accelerate flocculation,

the water is sometimes stirred with previously precipitated flocs, a process called 'suspended solids contact'.

Sedimentation. From the flocculating basin, the water flows to systems of clarification and sedimentation. In these basins, water moves slowly and the coagulated materials settle out. Water generally remains in these basins for several hours. The bottom of the basin has a scraper or rake which pushes settled material into a drain. The clarified water moves out from near the top of the basin. When the 'suspended solids contact' method of coagulation is used, a separate sedimentation tank is not needed, and water, after disinfection, goes directly to the filters.

Disinfection. Substances used for disinfection are chlorine, bromine, iodine and occasionally ozone and silver. By far the most commonly used of these is chlorine in gaseous form. Chlorine is a gas at normal temperature and pressure, but is usually obtained in pressurized containers in liquid form. The substance is released through a pressure valve as a gas and fed directly into the water. Sometimes sodium hypochlorite or calcium hypochlorite is used, but since their manufacture involves use of chloride as raw material, they are more expensive. It is thought that chlorine kills micro-organisms by destroying enzymes essential to life processes. Chlorine has strong oxidizing ability and must be handled with great care.

The quantity of chlorine desired in the water when it reaches the consumer varies some with conditions, but it is generally in the 0.2 to 1.0 milligramme per litre range. Since the chlorine concentration decreases with time, measurements are best taken after some contact time with the water and at various distribution points in the system. Chlorine concentration is easily tested using a simple colorimetric test. The safety of water for consumption is tested by examining for coliform (intestinal) bacteria, since pathogenic bacteria are at least as easily controlled as coliform bacteria, and the latter are quite easily tested for.

Filtration. The purpose of filtration is to remove any remaining particles of floc from the water. Several types of filters are used, but in municipal water treatment the most common is the sand filter. It is composed of a sand layer over layers of gravel. As water settles through the filter, particulates are collected in the sand layer and in the collecting layer of coagulated materials. When the filter becomes clogged, water is pumped up through the filter (called back-washing) and drained off. Filtration can then be resumed.

The importance of adequate sewage treatment

Among the effects of water pollutants are those caused by toxic heavy metals, sediments and organic materials with associated breakdown products (nutrients).

Heavy metals. Arsenic, chromium, copper, lead, zinc and mercury are illustrative of substances which are poisonous to various populations in natural ecosystems. Concen-

trations and effects vary and combinations of the substances may have different effects than the same substances occurring separately. For example, although certain game fish are not affected by 8 milligrammes per litre of zinc or 2 milligrammes per litre of copper, they are destroyed by a concentration of 0.1 milligramme per litre of the two combined. These materials are usually industrial wastes and their removal from water supplies, when known, is often quite complex.

Quantities and sources of water for human use

The amounts of water used by man vary greatly, according to the availability of natural waters and food materials as well as with the demands of increasing domesticity and industrialization. The average daily water requirement for a person is approximately 2 litres, some of which is obtained from fresh food sources. Table 27 shows the average volumes of water balance in humans. However, the 2 litres quoted as a survival amount will not necessarily ensure continuing health and well-being. Urine and sweat account for the greater loss of water from our bodies and the obligatory loss is dependent upon intake, except in cases of deprivation of water or extremely minimal intakes. The kidneys tend to regulate the release of water into the urine and the amount is restricted in conditions of inadequate intake, resulting in a concentrated urine. Environmental conditions in which the temperature and humidity levels are high stimulate a greater loss of water through the skin and less through the kidneys. (See Fig. 68 for water needs in arid conditions.)

Table 27
Average volumes of water balance in humans

Intake		Output	
Source	Volume (ml)	Source	Volume (ml)
Liquids	1 100	Urine	1 000
Foodstuffs	450	Perspiration, Respiration	900
Metabolic oxidation	450	Faeces	100
Total	2 000	Total	2 000

When examined closely, it is quite surprising how much water is used each day by an average household, but the data is even more significant when analysed historically. For example, in mediaeval Europe the average household consumed approximately 18 litres of water each day, whereas in modern households, living in more advanced technological societies, consume between 365 and 682 litres per day. This increased amount is explained more by the usage of water in bathing, the washing of habitations, clothes and kitchen utensils, domestic heating (with hot water or steam) cooking and the irrigation of gardens, car washing, etc., rather than by an increased physiological demand.

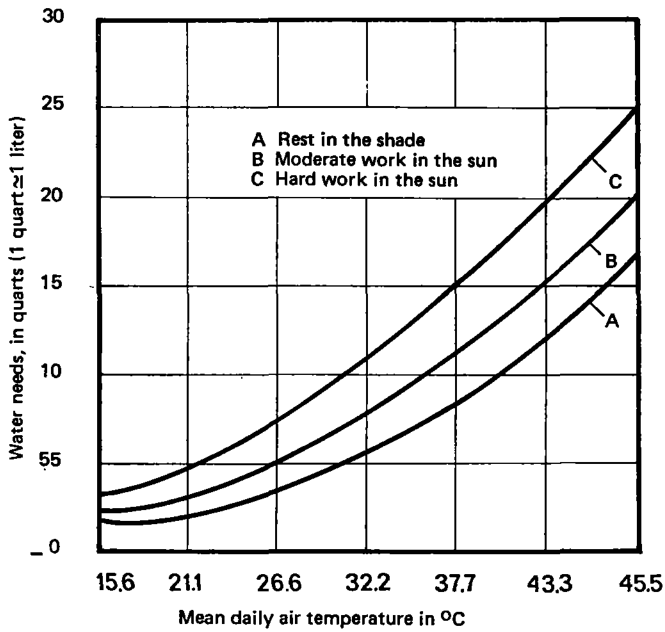


Fig. 68
Water needs for humans in the arid zone, according to the kind of activity. Needs corresponding to 16 hours in the shade (normal sedentary activities or rest) and to 8 hours of daily behaviour (of type A, B or C)

Source: Douglas H. K. Lee. 'The Problems of the Arid Zone'. Paris, Unesco, 1962. (Arid Zone Research, 18.)

When food chains, which are developed specifically for the consumption of foods by man are taken into account, the water demands are even more impressive. A food chain using wheat for bread-making requires about 1,135 litres of water per person, per day and the production of 1 kilogramme of beef requires 7,735–9,555 litres.

As industrialization continues, the very substantial additional amounts of water that will be consumed are predictable. For example, it takes 140 litres of water to make 1 kilogramme of sugar, about 1,800 litres of water to manufacture 1 kilogramme of rayon, about 18 barrels of water are required to refine one barrel of oil, and a paper mill will use an amount of water equivalent to a city of 50,000–150,000 people. Table 28 shows an estimate of the industrial water used by major manufacturing industries. In various industries, water performs such functions as lubrication, cooling, washing, power generation and waste disposal.

Man relies heavily upon the free waters of the hydrosphere for his ever-increasing demands. The total amount of such water resources approximates to 136×10^7 cubic kilometres. Of this amount 99.35 per cent is in the oceans, ice caps and glaciers and of the remaining 0.65 per cent, half is as sub-surface water. Hence human populations are dependent upon but a fraction of the great potential supply of the hydrosphere and this must be shared with the rest of the plant and animal world. Thus, it will be appreciated that stricter conservation measures are needed by all communities for the water which is available to us as rain (precipitation), in the rivers, streams and lakes and from underground sources.

In Asia and the Far East, the precipitation pattern varies from the dry desert areas of central Asia to extremely wet humid areas of Assam and the Province of

Table 28
Estimated industrial water consumption by major industries[1]

Industry	Production units	Water used (m^3/unit)
Steel	ton	250
Oil refining	cubic metre	18
Petroleum	cubic metre	9
Wood pulp	—	—
Sulphate	ton	240
Sulphite	ton	230
Soda	ton	320
Wood	ton	19
Paper	ton	150
Paper board	ton	57
Coke	ton	14
Soap	ton	2
Tanning	ton	67
Woollen fabrics	ton	580
Rayon	ton	1
Cane-sugar	ton	4
Ice	ton	1
Beer	ton	11

1. Most of this water returns to the river after use.

Source: S. T. Powell and H. E. Bacon. *Journal of the American Water Works Association*, August 1950.

Irian Jaya. Excluding the lower Indus Basin, the annual rainfall exceeds 1,000 mm over most of India and South-East Asia (in places exceeding even 3,000 mm). As elsewhere, the sources of water for various human use and consumption occasionally include direct precipitation and desalination, but are constituted mainly of groundwater and surface flow.

One-half of the world's population lives in Asia and the Far East. The region covers one-seventh of the global area and contains one-quarter of its available water and one-third of its potential hydro-power. The majority of countries in the region are in the developing stage where in water-power generation, mechanization of agriculture and industrialization are greatly emphasized. Most of the precipitation is seasonal and occurs in the monsoon or typhoon seasons and results in frequent destructive floods along the river courses and deltaic areas. On the other hand, there are regions in central Asia where rainfall is very meagre. Thus water is one of the biggest problems in Asia, at places the lack of it, while at others too much of it; however, both contribute one way or the other towards crop failure and other losses.

It is imperative that water conservation should be based on sound ecological principles taking into consideration all the interconnected and interrelated links influencing water supply. Proper attention to land use and better soil management including rotational irrigation can lead to better use of soil water and hence to soil conservation and diminished siltation downstream with the resultant increase in subsurface water. Better land-use practice on forest land and range land also help improve infiltration and stabilize run-off.

In many parts of Asia underground water is much in use. In places it is the only

source for all domestic supplies. As a result, most natural underground supplies have been depleted or are nearing depletion, thus causing the sinking of the land surface in land-locked areas or invasion of the aquifers by saline water from the sea in the coastal zones. Use of such underground reservoirs for the storage of excess run-off holds great potential. This can be achieved either by pumping in freshwater during heavy precipitation or by holding run-off where it can seep in through permeable soil.

However, water conservation for all human use cannot be met by efficient land use and soil management alone. Reliable metering and monitoring of water should be undertaken in various areas and an adequate levying of water charge should be directed at a level that would encourage greater efficiency in water use.

Water resources and their development

Life, as we know it, cannot exist without water. Water has been, throughout history, closely related to the rise and fall of civilizations. Human populations survive only where there is water adequate for all of their needs. One of the most important aspects of water resources development is the definition of purposes for which the water is required. These might be summarized as follows: increased agricultural development and productivity; provision of domestic and industrial water supplies; production of hydro-electric energy; protection against floods and sea-water intrusions; inland and coastal area navigation; improvement of fisheries and fish productivity; control of soil erosion; control of water quality.

Water resources development is a major involvement for every nation, both for its own well-being and for the benefit of its neighbouring states and the world. As a vital factor in our environment, water demands serious consideration as to its right use (conservation), hence good planning, proper management, rational legislation and enlightenment of all who use it, through the schooling programmes, must be brought into action and co-ordinated.

Indonesia with 1,940,000 million m^3 possesses the largest share of resources in the region followed by India with 1,508,000 million m^3 (total estimated run-off for Asia and the Far East being 6,700,000 million m^3). All this water will be wasted if not managed and used efficiently for one or more of the specific projects mentioned above. However, major water resource projects are costly and often strain the national financial resources to the utmost. In order to avoid any drag on the economy, the goals laid down should be achieved as planned and scheduled. This would be subject to efficient management and use of water. The benefits of water resources development may accrue either from increased agricultural production (for export or replacement of imported products) or hydro-electrical energy (reduce demands for imported fuels, etc.) with an over-all rise in living standard and self-sufficiency.

For any water resources project to be successful there should be good planning, proper management and all the necessary support from the water users.

In developing countries, the planning, construction and operation of irrigation and drainage projects are often similar to those practised in the developed countries

regardless of differences in conditions. However, subsistence farmers with small farms and using traditional farm practices present managerial problems not encountered in the developed countries. In such a case it would be necessary to have water-users organizations affiliated to some government agency responsible for improving farm inputs such as seeds, fertilizers and insecticides. Thus co-operative marketing and purchasing can be arranged. Such a water-users organization set up in Taiwan some years ago has proved to be very effective.

In many Asian countries, projects have been developed in such a way as to channel water to areas of land occupying 50 to 100 hectares. It was then left to each individual farmer to devise ways and means of transporting the water to his farm. Fortunately attempts are now being made to provide water control to each farm separately in the newly established project areas.

The long dry season associated with the monsoons accounts for the rivers in Asia running dry. Rivers which have their sources in the mountainous hinterland (Mekong and Brahmaputra) never run completely dry even in the dry season. There is thus need for irrigation in both the wet and dry seasons.

Under the Indicative World Plan (IWP), it is intended to increase the irrigated area from 64.1 million ha to 92.3 million ha in the next twenty years and at the same time to reduce the non-irrigated area from 234.0 million to 224.3 million ha.

The irrigated harvested area will also increase from 76.2 million ha to 136.2 million ha. There will be a slight increase in the non-irrigated harvested area from 221.9 million ha to 230.2 million ha.

On both irrigated and non-irrigated areas there will be a corresponding increase in cropping intensity from 119.0 to 147.7 per cent in irrigated land and from 95.0 to 102.5 per cent on non-irrigated land as shown in Table 29.

Table 29
Planned increase in irrigated area and irrigated harvested area (million ha)

Year	1970	1975	1980	1985	1990
Cultivated area (total)	298.1	302.9	307.5	312.1	316.6
Irrigated	64.1	69.0	76.0	83.4	92.3
Non-irrigated	234.0	233.9	231.5	228.7	224.3
Harvested area (total)	298.1	314.1	331.4	348.5	366.4
Irrigated	76.2	87.6	102.8	118.7	136.2
Non-irrigated	211.9	226.5	228.6	229.8	230.2
Cropping intensity (percentage)	100.0	103.5	107.8	111.7	115.9
Irrigated	119.0	127.0	135.0	142.0	147.7
Non-irrigated	95.0	96.8	98.9	100.5	102.5

Source: Water Resources Series. ECAFE (U.N., New York), 1971, 40, p. 34.

India will have to expand its irrigated area from 32.6 million ha to 46.5 million ha and its irrigated harvested area from 42.9 million ha to 77.1 million ha.

Pakistan will likewise need to expand its irrigated land from 13.0 million ha to 20.4 million ha and its harvested area from 15.6 million ha to 28.2 million ha. Other countries too will have to expand at comparable rates to satisfy their demand for food.

Similarly in the field of agriculture in India the production of grains has nearly doubled since independence. This has been made possible through technological advances like the use of high-yielding seed varieties, increased use of chemical fertilizers and insecticides coupled with improved farm practices.

The high yielding varieties of seeds require much more watering and, to achieve optimum production of food grains, it is imperative that there be a more scientific approach to the management of water. Moreover, it must be borne in mind that crops like sugar-cane and rice require quantities of water whereas crops like maize and wheat need much less. Till now not much attention has been focused on the problem of water supply for agriculture.

Irrigation water should be managed in such a way as to minimize wastage of water, soil and plant nutrients and maximizing crop yields. If water is applied in too great quantity it will cause drainage congestion, if insufficiently applied it will not meet leaching or crop requirements and if the rate of application of water is more than the optimum there is bound to be erosion.

In irrigation practices, the concept now followed is to provide adequate water from surface resources to the farmers for one crop and leaving it to them to make use of groundwater for the second and subsequent croppings. Irrigation water is supplied on the basis of harvested area except of course when the area is small as in Maharashtra. Here water is provided on a volumetric basis for sugar-cane cultivation. This latter method prevent undue wastage of water.

7 Soils

Introduction

Wherever human populations have been able to develop natural soils in terms of improved fertility, farming and land use, settlements and civilizations have been well established and extended. Soils are an intrinsic factor in the development of natural ecosystems, including, more particularly, those involving human populations. Thus, this chapter provides a survey of soil composition, the factors affecting soil, soil classification and human interactions with the soil. An understanding of all of the factors relating to soils is important, in order that the roles of soils, as environmental influences on human populations, can be appreciated. Human interactions are, therefore, considered primarily in terms of agriculture and, incidentally, in relation to urban development.

Definition of soil

Soils are generally considered to be sediments, deposited by natural accumulation of appropriate materials from the bed-rock, over which such soils form a mantle; alternatively, by the action of wind, water and large-scale agriculture. The sedimentary nature of soils is, nevertheless, extremely variable in texture, composition and fertility. Such characteristics create a difficulty in clear definition, hence soils are more often identified either in terms of origin or fertility in relation to the flora, either indigenous or agricultural. The scientific study of soil, known as pedology, developed since the beginning of this century, with particular reference to temperate zones and less so in intertropical regions. Tropical soil studies have yielded a wealth of valuable data, particularly from the Asia region, assisting in an improved understanding of the relationship between soil development, as a natural phenomenon, and the impact of human populations on soils.

Composition of soil

Soil results as a product of two operations: the decomposition of rock, the decay of organisms. Natural soil originates as unconsolidated mineral material. It is later mixed with varying amounts of organic materials (humus) arising from the decay of organisms that have lived either within the developing soil or on its surface.

Soil is a complex system that provides mineral elements, water and soil air for all kinds of plants growing in or on it.

The characteristics of soil that will be discussed here are: soil types, texture, soil water, soil air, soil organisms and organic matter (or humus).

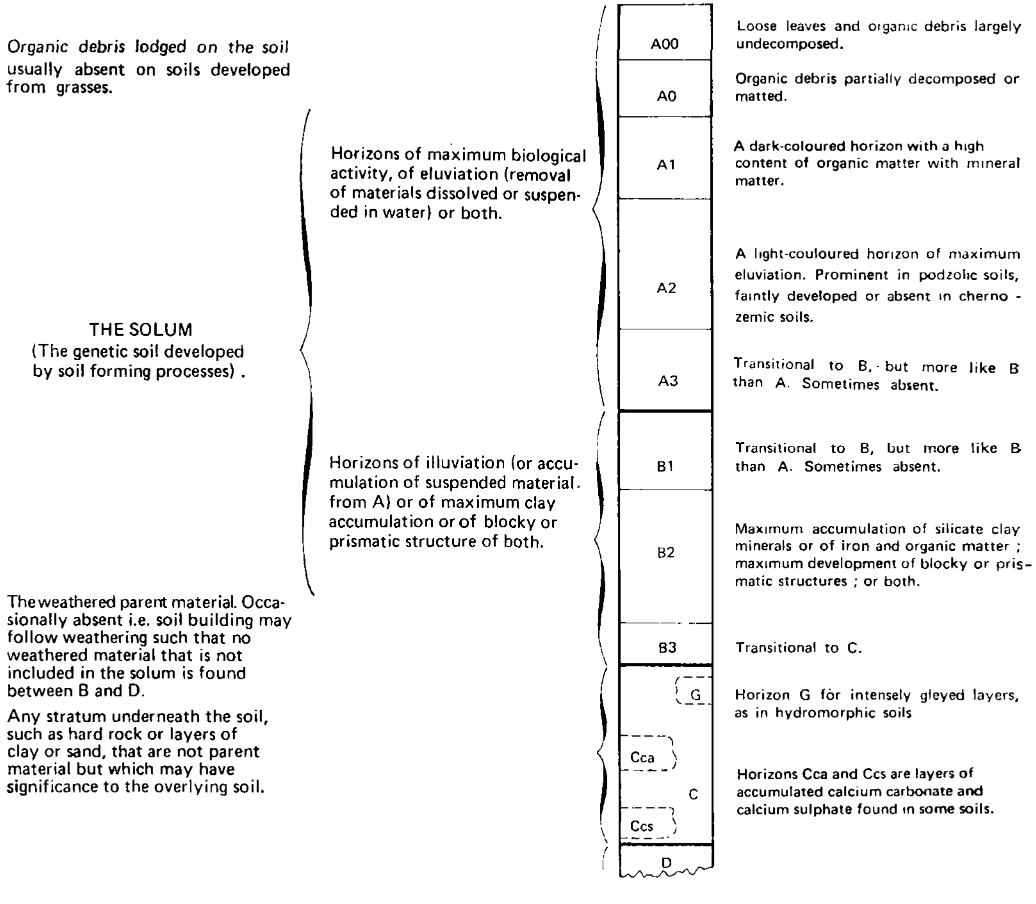


Fig. 69
A hypothetical soil profile having all principal horizons. It will be noted that horizon B may or may not have any accumulation of clay.

Horizons C_{ca} usually appear between B_3 and C. The G may appear directly beneath the A (Courtesy U.S. Soil Conservation Service)

Source: R. V. Tamhane; D. P. Motiramani; Y. P. Bali; Roy L. Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia.* New Delhi, Prentice-Hall of India (private) Ltd., 1964, p. 91.

Soil types

Although soil covers the earth's crust (lithosphere) in an almost continuous mantle, it differs greatly in character as well as in usefulness to human populations, from one region to another. Such character differences are noticeable even around a village or within a valley as well as on a piece of farm land. Soil scientists and others have attempted to classify the different soil types, based on the origin and properties of each soil. Thus modern soil classification includes an appreciation of the distinctive layers or horizons, formed above the consolidated rock strata underlying the soil mantle, and the sequence of the horizons at a given place. The horizon arrangement is determined by the interaction of climate, vegetation, topography and time together with the geochemistry of any contributing rocks and minerals. Fig. 69 shows a hypothetical profile having all principal horizons, while two contrasting Indian soil profiles are shown in Table 30.

Table 30. Contrasting soil profiles

Soil profile from a village in New Delhi		Soil profile from a village in Madhya Pradesh	
Average annual rainfall is 66 cm.		Average annual rainfall is 132 cm.	
Profile description			
0–10 cm.	Light yellowish brown sandy loam, weakly granular with some weak lamination, loose and soft when dry, no effervescence with dilute HCl, pH 7.7.	0–23 cm.	Reddish brown, gravelly silt loam, weak medium sub-angular blocky, hard when dry, many iron concretions, rapidly permeable, pH 6.8.
10–36 cm.	Dark yellowish brown sandy loam, somewhat crumby to sub-angular blocky, very hard when dry, no effervescence, pH 8.5.	23–58 cm.	Brown, gravelly clay loam, structure indistinct, due to abundance of iron concretions, fairly permeable, pH 6.7.
36–76 cm.	Dark yellowish brown clay loam, sub-angular blocky, very hard when dry, slight localized effervescence on small calcareous concretions, pH 8.9.	58–89 cm.	Reddish-brown, gravelly clay loam, mixed with iron concretions, fairly permeable, pH 5.9.
76–127 cm.	Yellowish brown clay loam, blocky, very hard when dry, calcareous concretions increase in size and quantity, localized effervescence, pH 9.2.	89–109 cm.	Red compact layer of iron concretions.
127–188 cm.	Yellowish brown clay loam, same as above, pH 9.5.		

Source: Tamhane, Motiramani, Bali and Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia*. New Delhi, Prentice-Hall of India (Private) Ltd., 1970.

We owe our present understanding of soil classification to the pioneer researches of K. D. Elinka and his colleagues working at the Leningrad Agricultural Institute. Table 31 gives a simple classification of tropical soils. Mention must also be made of the agricultural potential of soil types related to pedogenesis (soil origin).

Table 31
Simplified classification of tropical soils

Azonal soils (little profile development shown).
1. Fresh alluvium, lithosols and regosols.

Intrazonal soils (effects of parent rock or drainage factors or shortness of time dominant over those of climate and vegetation).
2. Dark cracking clayey soils (grumusols and non-grumusols).
3. Allophane soils formed from volcanic ash and agglomerate.
4. Calcimorphic soils.
5. Hydromorphic soils.
6. Halomorphic soils.

Zonal soils (effects of climate and vegetation acting over long periods dominant over those of parent-rock and drainage factors).
7. Pedocals.
8. Pedalfers: latosols; podsols (not including groundwater podsols); fersiallitic soils (including some ferruginous siallitic soils).

Source: C. C. Webster; P. N. Wilson. *Agriculture in the Tropics.* London, Longmans, 1968, p. 30.

In brief, tropical soils have been divided into three broad groups—zonal, azonal and intrazonal. Zonal soils are those which have developed under the influence of climate and vegetation. Azonal soils or juvenile soils are still in the formative stage, i.e. there is just the beginnings of stratification. The intrazonal soils are well-developed soils but with their classification based on relief, parent material or time.

Fig. 70 shows the primary soil types of the world and the most important soil groups of Asia are described. (See also Figs. 71 and 72.)

Alluvial soils

Alluvial soils do not exhibit prominent horizon development as they are developed from water-deposited sediment. These soils are mostly acidic, greyish in colour and are poorly drained. In Asia these soils are formed in river basins on flood terraces, flood plains and deltas and are a familiar sight in the Indo-Gangetic Plains of India.

Regosols

The regosols are developed from unconsolidated materials like volcanic ash, marls and sand dunes and thus do not have a well-developed profile. Regosols are found in the coastal areas of Asia as well as in the desert areas of India and Pakistan. They are also common in volcanic ash in most of the islands of Indonesia and the Philippines.

Lithosols

Lithosols are shallow soils found on steep slopes where there has been limited weathering of the underlying rocks.

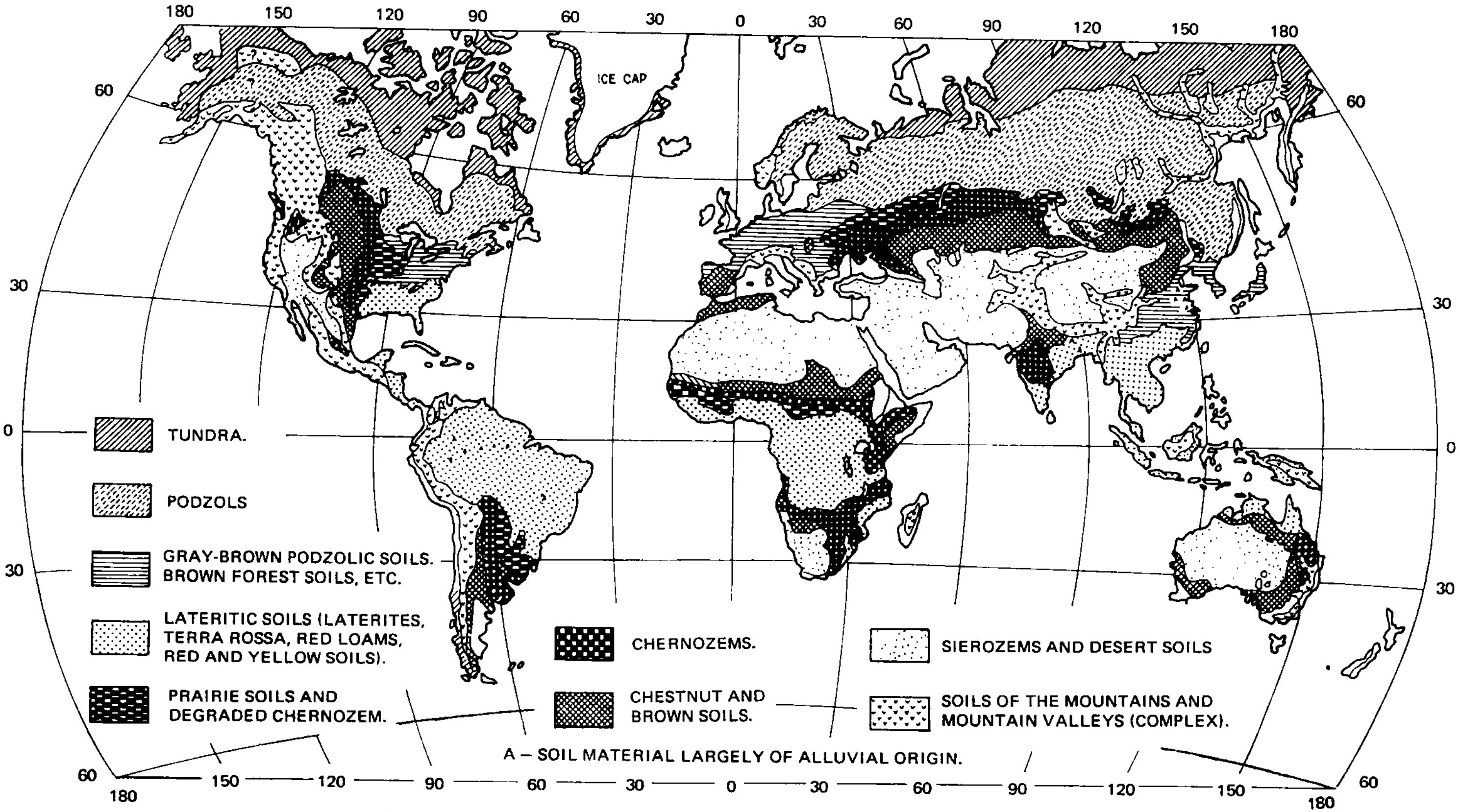


Fig. 70
Primary soil types of the world

Source: Eugene P. Odum. *Fundamentals of Ecology*, 3rd ed. Philadelphia, W. B. Saunders Co., 1971.

Note. Not only north-south differences are evident, but also east-west differences which are related to rainfall. Podzolic and lateritic soils of humid regions are often known as 'pedalfers' because of the accumulation of iron and aluminium in the B horizons, while the chernozems and other soils of more arid regions are 'pedocals', because of the accumulation of calcium. In recent years a more uniform and descriptive set of names has been proposed for zonal soil types, each name having the same root-'sol'. Thus, northern podsol becomes spodosol (=ashy), the temperate gray-brown podzolic type becomes alfisol (Al-Fe, referring to B_2 material accumulation), the prairie soil becomes aridisol, and tropical, lateritic soil becomes oxisol (=oxidized). Map from U.S. Department of Agriculture Yearbook, 1938.

Andosols

Andosols found commonly under forests are acidic in reaction and are well supplied with nutrients. In Indonesia and the Philippines they have been detected on volcanic ash under humid conditions in mountainous areas.

Podzol soils

Podzol soils which are acidic in reaction and coarse in texture are also found under forests. In Asia, where there is an accumulation of much organic matter, humus podzols are of frequent occurrence.

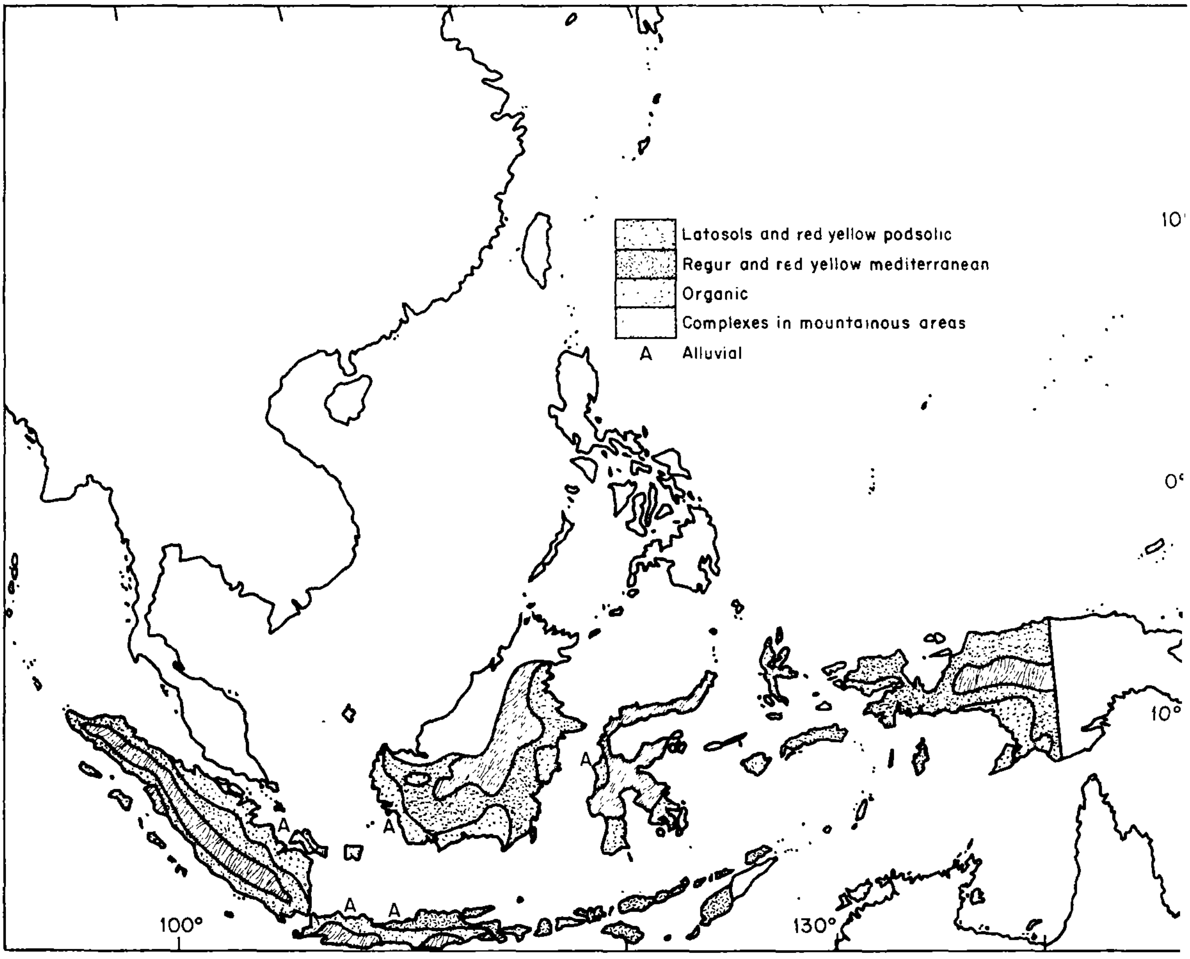


Fig. 71
Generalized soil map of Indonesia
Source: Tamhane, Motiramani, Bali and Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia*. New Delhi, Prentice-Hall of India (Private) Ltd., 1970, p. 117.

Latosols

Latosols include all the soils which are deeply weathered and strongly leached and which show no horizon differentiation. They include all the lateritic soils ranging in colour from red to brown to yellow. Latosols which are acidic in reaction occur extensively in Taiwan and India.

Grumusols

Grumusols include the dark clay soils which have been given different names in different Asian countries. For example the grumusols are often referred to as the

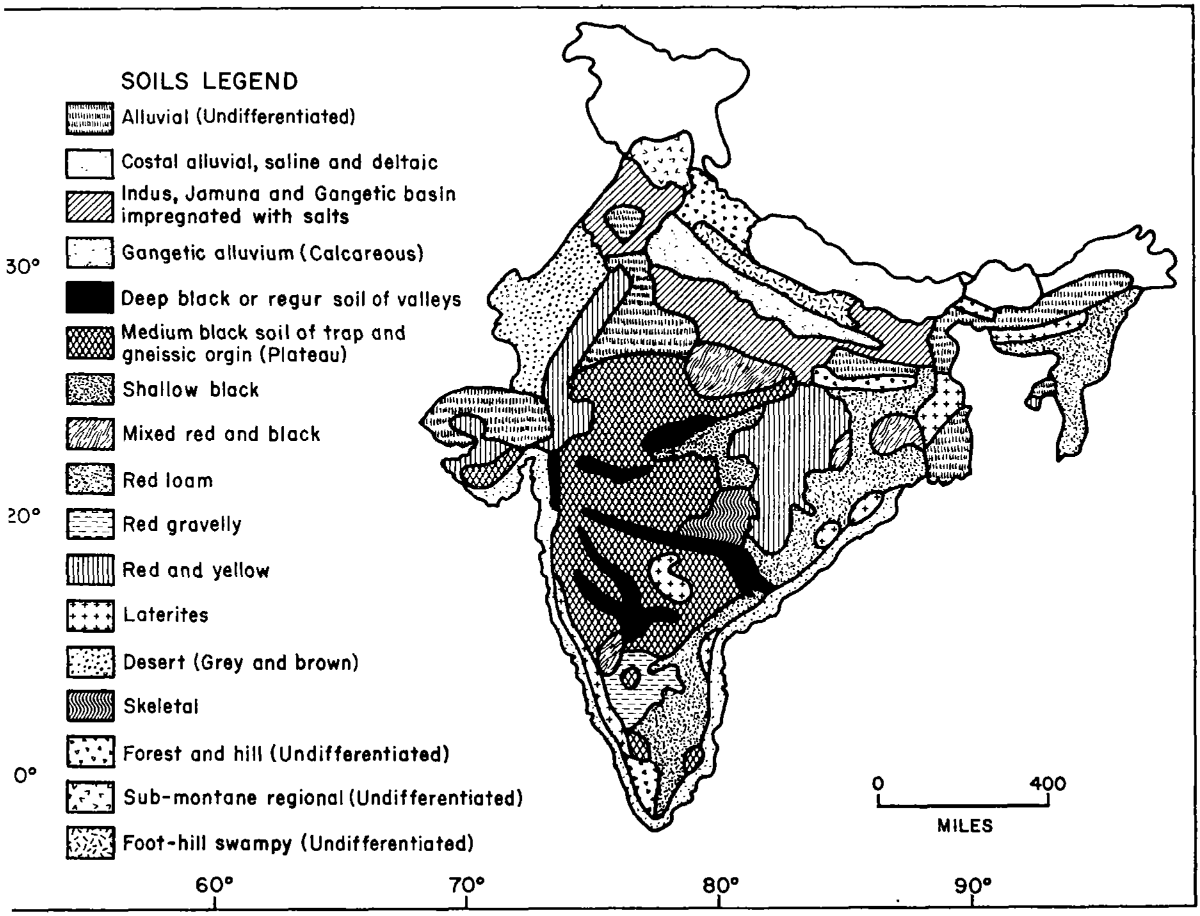


Fig. 72
A generalized soil map of India

Source: Tamhane, Motiramani, Bali and Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia*. New Delhi, Prentice-Hall of India (Private) Ltd., 1970, p. 118.

black cotton soils in India, the regurs in India, Khmer Republic and Indonesia and margalites and black earth in Indonesia. Grumusols are generally poor in organic matter but rich in clay content and the latter is a uniform feature throughout the profile. As these soils exhibit plastic properties when wet they tend to produce deep cracks with changes in moisture content.

Grey hydromorphic soils

These soils are commonly referred to as low humid gleys. They exhibit a brownish grey to livid grey mottling in the B horizon while the surface horizon is very thin. These hydromorphic soils are not uniform in distribution but appear in patches in most of the tropical Asian countries.

Land areas without soil

A major problem, found in many regions of the world is the lack of any soil cover whatsoever. In tropical latitudes, the soil-less areas coincide with the dry subsiding air movements and horizontally diverging wind systems of the subtropical anticyclones. (Incidentally, these are also found in the polar regions.) The nature of the topography, when erosion by rain and wind removes any loose, weathered material from the earth's surface, also contributes to soil-less land areas. Soil-less areas are found in the tropics in general and in Asia in particular. The data given in Table 32 is based upon bare rock surfaces, desert areas, lake and large river surfaces.

Table 32
Areas of the world without soil

Continent	Source	10^6 km^2	Percentage of continent
Africa	d'Hoore (1965)	8.46	28.89
Antarctica	Gracanin (1967)	14.00	100.00
Australia	Gracanin (1967)	1.28	14.27
North America	Gracanin (1967)	2.69	11.06
South America	Gracanin (1967)	0.29	1.61
Europe/Asia	Gracanin (1967)	3.38	6.34
Total for world		30.10 10^6 km^2	
Percentage total land areas		20.22	

Soil texture

Soil texture is determined generally by the soil particles, produced through the weathering of parent materials derived from the underlying rocks, through transportation from other areas. The size of particles varies from boulders and gravel down to sand and sub-microscopic particles of clay. Large particle size allows comparatively rapid permeation of water and facilitates diffusion of soil gases. Soils of small particle size retain water more effectively and allow greater rise of water from an underlying water table due to capillarity. The relative size of soil particles, based upon the International and Mohr's Systems, is given in Table 33. (See also Tables 34 and 35.)

Table 33
The relative size of soil particles by mechanical analysis

International system[1]		Morh's 10-fraction[2] system	
Soil separate	Diameter range (mm)	Soil separate	Diameter range (mm)
1. Coarse sand	2.00–0.20	1. Very coarse sand	2.00–1.00
2. Fine sand	0.20–0.02	2. Coarse sand	1.00–0.50
3. Silt	0.02–0.002	3. Medium sand	0.20–0.10
4. Clay	Below 0.002	4. Fine sand	0.20–0.10
		5. Very fine sand	0.10–0.05
		6. Coarse silt	0.05–0.02
		7. Silt	0.02–0.005
		8. Fine silt	0.005–0.002
		9. Clay	0.002–0.0005
		10. Colloidal clay	Below 0.0005

1. International Society of Soil Science Systems.
2. Mohr's System, applied to tropical soils.

Table 34. Categories of particles and soil textures

Fractions	Clay soils	Clay silts	Sandy silts	Loose soils	Moist humic sands	Moist humic clays	Calcareous clays	Calcareous sands
Clay 0.002 mm	25%	20	5–10	5–10	1–5	20	20	5
Silt 0.002–0.02 mm	25–50%	25–40	30–40	10–15	10–15	10–20	10–20	5–10
Fine sand 0.02–0.2 mm	20–25%	20–30	20–40	15–30	15–30	15–30	15–30	15–30
Coarse sand 0.2–2 mm	5%	5	20	30–50	30	5–15	5–15	30–40
Calcium carbonate	1%	1	1	1–5	1–3	1–3	5–30	5–30
Organic matter	3%	3	3	3–5	10	10	3	3

Source: Unesco. *La plante et le sol. Vol. 1. Les sols.* Projet-pilote pour l'enseignement de la Biologie en Afrique et à Madagascar. Paris, AUDECAM, 1971, 84 p.

Table 35. Texture and type of soil

Different soil textures	Types of soil	Cultivation possibilities
Very fine clay texture, very rich in mineral colloids	Heavy soils	Difficult to cultivate
Coarse or sandy texture	Light, loose soils, often dry	Easy to cultivate
Balanced texture	Soils which are a mixture of sand, clay and silt, containing not more than 35 per cent silt, which give a loose texture from the agricultural point of view and which have good physical properties because of the balance between the coarse elements and the mineral colloids	Easy to cultivate
Silt texture	Soils rich in fine and coarse silts containing insufficient colloids, giving rise from the agricultural point of view, to so-called earth pan, having poor physical properties and permeability, and bad aeration	

It is possible to see a connexion between the texture of the soils studied and the agriculture carried out with these soils: rice prefers clay soils; ground nuts prefer sandy soils or clay-sand; the coco-nut requires soils of a clay-silt-sand composition; the oil palm is to be found in soils containing not more than 45 per cent of mineral colloids.

The porosity of a soil is determined by its total pore space, i.e. the unit volume of soil not occupied by solid materials. Porosity varies from 35 to 50 per cent in dry soils. Generally, the finer the texture, the greater the decrease in size of individual pore spaces, whereas the total pore space increases. Structural changes and variation in organic matter content also affect pore size and the degree of porosity.

Permeability in soils is regulated by the porosity as well as by the degree of aggregation of mineral particles and it is to be emphasized that in clay soils and highly enriched humus soils that the colloidal properties can impede normal permeation of water and hence control the general permeability.

In clay soils the clumping of soil particles into aggregates is an important structural feature. The presence of the aggregates facilitates permeability of the soil to water, air and plant roots while still promoting high water and nutrient-holding capacity. When the clay is associated primarily with univalent ions (such as sodium), soil particles are dispersed into a single grain structure which is much less suitable for agriculture. If hydrogen ions occur in excess, a good aggregate or 'crumbs' structure may occur, though it is less stable than when based on the activity of calcium ions.

Soil water

The water in the soil contains a variety of mineral (inorganic) and organic substances in solution. The composition of water in soils varies, depending upon the nature of its solution. High mineral content increases the salinity of soil water and makes the soil unsuitable for plant growth. Conversely, when the minerals in solution are very low in quantity and quality then there is a general deficiency in the nutritional status of the soil. The osmotic potential of the soil solution follows the general properties of an osmotic system. If the osmotic potential is very high as in distinctively salty (saline) soils, then this is a limiting factor of agricultural significance, in so far as such a soil will support only plants that are physiologically adapted to saline soil conditions. There are many such forms of plants occurring in natural vegetation and these are called *halophytes*.

Salinity is an agricultural problem in arid and semi-arid regions where there is inadequate rainfall to leach mineral matter downward in the soil, where the water table is close to the surface and evaporation exceeds rainfall and where there is irrigation along with high evaporation and poor drainage. Alkaline soils are high in sodium or potassium ions, resulting in poor drainage and poor aeration of the soil. Some of the major cations of soil solutions are Ca^{++}, Mg^{++}, K^{+}, Na^{+}, Al^{+++} and Fe^{+++}. Some major anions are HCO_3^-, $H_2PO_4^-$, Cl^-, NO_3^- and SiO_3^{--}.

A group of chemical substances, known as essential elements, play a vital role in the general fertility of soil and more particularly in soils that are intensively cultivated and subject to agricultural usage. In the mineral (inorganic) fraction of soils, silicon predominates. Silicon occurs as the oxide or as silicates of aluminium and oxygen. Hydroxides and carbonates form the natural compounds of iron and the basic elements, and these may be further combined with silicon and aluminium. In general, plant nutritionists consider sixteen elements essential to the normal, heavy growth of plants. A further three, namely sodium, iodine and cobalt, are needed by animals. Many soil elements are toxic to living things, if available in excessive concentrations. Table 36 summarizes the essential elements, assimilated by plants and animals, found in soils.

The mineral elements for normal plant growth may be classified in two categories: (a) those forming 99 per cent of living plant material, such as carbon, oxygen, hydrogen,

Table 36
Macro- and micro-nutrients found in soils

Essential elements obtained from air and water	Macro-nutrients	Micro-nutrients	Elements essential to animals	Toxic elements
Hydrogen	Nitrogen	Chlorine	Sodium	Selenium
Oxygen	Phosphorus	Manganese	Iodine	Arsenic
Carbon	Potassium	Copper	Cobalt	Molybdenum
	Sulphur	Iron		Fluorine
	Calcium	Zinc		Aluminium
	Magnesium	Boron		Nickel,
		Molybdenum		etc.

phosphorus, nitrogen, potassium, sulphur, magnesium and calcium; (b) those contributing to the essential metabolism and development of the plant, such as iron, zinc, manganese, boron, copper, molybdenum.

While zinc, copper and molybdenum function as catalysts in the biosynthesis of many organic compounds, nitrogen, phosphorus and sulphur are the basic constituents of proteins and nucleic acids in all living matter.

Potassium, calcium and magnesium function as cations in many ways, particularly in maintaining the cation-anion equilibrium in plant tissues.

Soil air

Soil air occurs in the spaces between the soil particles (interstices) which are the result of irregularities in the shape and size of the particles. Large spaces facilitate drainage and aeration while small spaces enhance water retention. Thus, a range of particle sizes is generally desirable in a soil. Pore space varies from about 30 per cent, by volume, for sandy soils, to approximately 60 per cent, by volume, for clays. There is no direct relationship, however, between the general texture and pore space. Working the soil, as in ploughing, tends to increase the pore space while heavy top or channel irrigation decreases the pore space. The majority of crop plants need a discrete balance between soil water and soil air in order to grow successfully.

The composition of soil air is quite different from that of the atmosphere, varying with such factors as the degree of humus formation and microbial activity, the climate pattern and drainage. The change in the composition of the soil mineral status also plays a part in regulating the nature of the soil air. Carbon dioxide concentrations are relatively high and result from the respiratory activities of the micro-flora and micro-fauna as well as of plant roots. If CO_2 exceeds 5 per cent, by volume, then plant root systems suffer adversely, due to the toxic effect of the gas. Oxygen normally enters the plant root system in aqueous solution, but there are a number of plants with their root anatomy adapted to O_2 concentrations in the soil air, these being able to increase the rate of absorption of the gas.

Organisms in the soil

The most abundant and widespread organisms in soil are various forms of bacteria, mainly heterotrophs, but with significant species of autotrophs and symbiotic forms in many soils. In relation to soil fertility a major effect of bacterial action is to provide nitrogen in chemical combination, usable by most green plants. Other bacteria play a vital role in sulphur metabolism while others are deployed in cellulose decomposition. This last function is important in humus formation, which is discussed later in this chapter. Fungi are abundant in acidic soils and function in the decomposition of organic matter.

The soil fauna lives either on other organisms or on organic matter in the soil;

they contribute in humus formation. Among the most important soil dwellers are the protozoa, various insects, including termites, ants and beetles, earth-worms, mites, nematodes and springtails millipedes and burrowing rodents. In tropical soils, the termites (Isoptera) live on wood and other plant matter, excavating and taking large quantities of soil for their galleries and mounds. They partly digest significant amounts of organic material, adding combined nitrogen to the soil in finely divided particles. Termites mix, aerate and loosen the soil, thus affecting the texture; soil turnover is achieved by their activities in much the same way as with earth-worm behaviour, typical of temperate soils. It is to be noted that the destructive effects of termites regarding crop plants must be considered against their valuable contributive features when resorting to control measures and the application of pesticides.

Earth-worms are quite common in some areas of the wet tropics, showing themselves on the soil surface during the rainy season. These animals tend to live well down in tropical soils, away from the drier levels, as they are much dependent for survival on moist environments. In savannah grassland soils, with a high percentage of organic matter and adequate rainfall, the earth-worm plays an important role in soil fertility. In the dry tropics, earth-worms are seldom seen and may not exist in the soils of these regions.

Soil-dwelling nematodes may attack plants (e.g. *Heterodera* sp.), causing diseased conditions, or they may be parasitic in mammals, including wild animals, domestic livestock and man (e.g. *Filaria* sp., *Ascaris* sp.). They are quite numerous in wet, tropical and subtropical soils, particularly those rich in organic matter. In the soils of the dry tropics and subtropics, nematodes are seldom found, probably because of the arid conditions in which they are unable to survive.

Organic matter and humus

The partial decomposition of soil organisms and especially plant residues provides the organic matter of the soil.

One of the functions of soil micro-organisms is the decomposition of organic matter that arises from leaf-fall, plant residues, dead roots, stubble, green manures and animal excrement. The carbohydrates and similar substances serve as a source of energy for soil bacteria, fungi and actinomycetes and are the chemical origins of the large amounts of carbon dioxide found in soil air. Other organic compounds more resistant to decay, remain in the soil for lengthy periods and are important components in humus formation. Alternating periods of relative dryness and moist conditions contribute towards the regulated breakdown and transformation of organic matter into humus. Such periods also inhibit the occurrence and activity of soil micro-organisms. Muddy conditions accelerate the decomposition process and favour the hyper-activity of soil micro-organisms. However, the rate of humus formation is at its optimum under constantly damp conditions, other factors being equal. In sandy soils the inorganic components exert little influence on the decomposition of organic matter, nevertheless, because of the physical nature of sandy soils, humus formation

is very dependent upon the status of the micro-organisms, water and soil air. Organisms associated with each other in humus formation are the ammonifiers, the denitrifiers, fungi, starch decomposers and protein decomposers.

In many tropical and subtropical soils, conditions are excellent for decomposition by micro-organisms. As a result, organic matter is rapidly decomposed and very little humus accumulates even though large amounts of organic matter may be supplied to the soil by the surface vegetation. (See Fig. 73 for organic matter relationships under forest.)

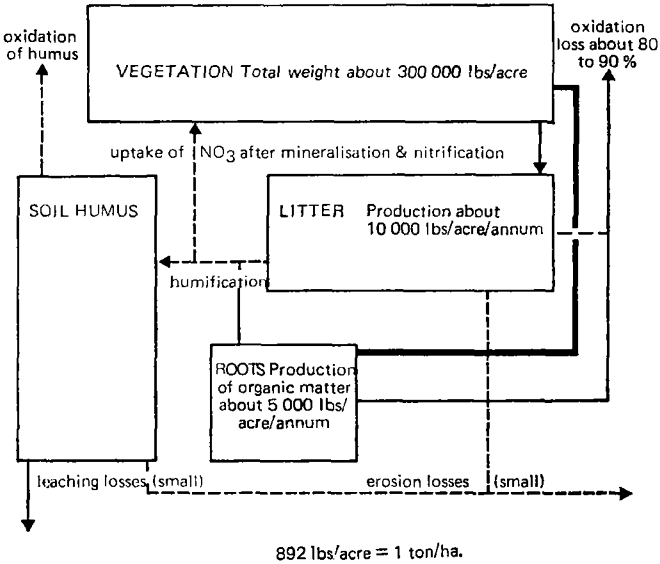


Fig. 73
Organic matter relationships under forest
Source: Thomas and Whittington. *Environment and Land Use in Africa*. London, Methuen, 1969.

Factors affecting soil formation

Soil formation is a complex process which cannot be discussed with any adequacy here, but it suffices to say that soil is the ultimate result of the actions and interactions of five soil-forming factors: parent materials, time, relief, climate and biota including man. Any one of the five factors may influence soil formation to a greater or lesser degree than the other remaining factors. One method of expressing the relationships of the contributory factors to each other and to soil properties is as follows:

$$S = f(cl, b, r, p, t \ldots)$$

where S = any soil property e.g. soil content, f = function of or dependent upon, cl = climate, b = biota (vegetation, organisms, man), r = relief (topography), p = parent material and t = time (age).

It is evident from the above equation that any evident soil property is a function of the other soil-forming factors.

Another useful way of appreciating soil factors is to categorize them into two

groups: passive factors and active factors. Passive factors include parent material, relief and time, while the active factors are climate and biota.

Parent material was, at one time, emphasized as the primary basis for differentiating soils. The correspondance of plant distribution in Europe with distribution patterns of substrate rock was a significant factor in this appraisal. It was then discovered that soil development processes are similar, regardless of parent materials, under uniform climate. Hence, the role of climate has been emphasized more strongly in recent years. Nevertheless parent material continues to be recognized as important in soil formation, since the composition of the rock being disintegrated (or transported into the area) establishes the salient mineral characteristics of the soil which develops.

The role of time is fairly clear at the extremes. In very old soil, little or no weatherable rock may be left, while soil formation may have barely begun in a young soil. Other factors may cause extensive modifications of the effects of time. For example, in humid tropics soil formation is almost continuous, while in an arid region, change is likely to be much slower.

Topography affects soil erosion, but other factors are correspondingly influential concurrently. The gradient of a slope affects the rate and amount of run-off, but the presence of vegetation and its root structures may also be important in determining the rate and amount of run-off.

Climatic factors and the biota in soil development are of major significance to human populations. These are discussed separately below.

Climatic factors

Climate influences the nature of natural vegetation, soil characteristics, the types of farming practised and to a greater extent it determines the crops that can be grown in any region. Thus, there is a relationship between the diverse types of tropical and subtropical climates, the soils and the agricultural potentials of any given land area. More detailed information is to be found in the 'Further Reading' section at the end of Part II.

Tropical Asia lies mainly between 10° South and 37° North latitudes and between 62° and 130° East longitudes. Thus most of tropical Asia is situated in the tropics but the more northerly parts of India, Burma and Pakistan belong to the subtropics.

Near the equator, the temperature is generally quite warm and constant (Fig. 74). In the tropical rain forests, high humidity and rainfall also prevail, resulting in rapid and continuous weathering by chemical action and leaching. Decomposer activity is also likely to be high under these conditions as the velocity of decomposition is approximately doubled for every 10° rise in temperature. In arid regions, precipitation and biological and chemical activity are likely to be much lower and the rate of weathering comparatively low.

Tropical Asian soils have been mainly developed from granites, gneisses, basalts, sandstones, limestones, shales, volcanic ash, alluvial deposits, littoral deposits and raised coral reefs. Soils developed from igneous and metamorphic rocks have bright yellow or dark red colours because of the intense oxidation and hydration. The high

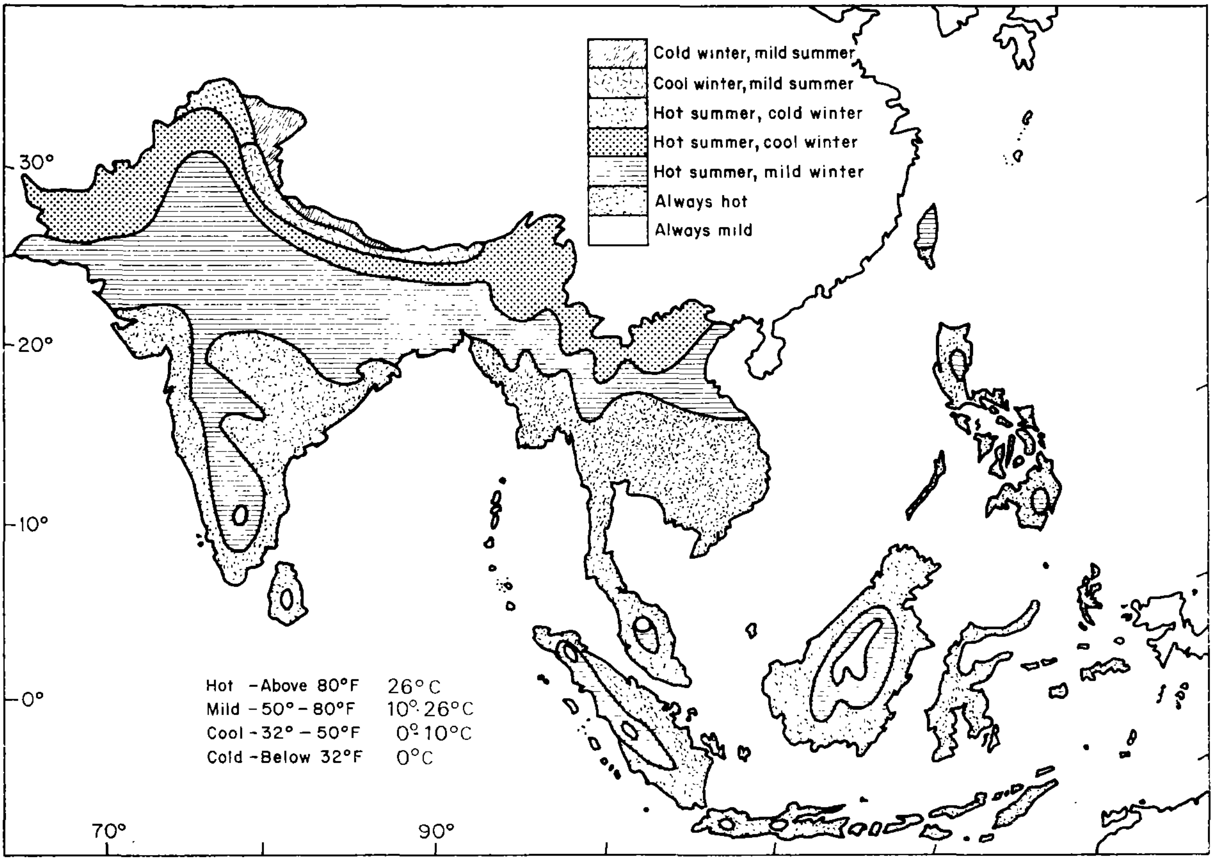


Fig. 74
Surface temperatures
during summer in tropical Asia
Source: Tamhane, Motiramani, Bali and Donahue. *Soils: Their Chemistry and Fertility in Tropical Asia*. New Delhi, Prentice-Hall of India (Private) Ltd., 1970, p. 98.

temperature in the tropics is not favourable for the accumulation of organic matter in the soil, since microbial activity is high resulting in rapid decomposition of organic matter. Hence, tropical soils are often low in organic content.

Biotic factors

Soil formation is greatly influenced by the various biotic factors, for example, vegetation, micro-organisms, animals and man. The following list is indicative of summary statements concerning biotic factors and soil.

1. Vegetation is a major source of soil organic matter, though animal bodies and animal wastes also contribute.
2. Bacteria and fungi decompose organic matter in the soil.

3. Soil animals such as termites, ants and earth-worms mix and aerate the soil.
4. Vegetation breaks the impact of rain on the soil, aiding penetration of water and reducing erosion.
5. Leaf litter (decomposing vegetation on the surface of the soil) breaks the impact of rain on the soil, aiding penetration of water and reducing soil erosion.
6. Litter reduces evaporation from the soil and assists in heat conservation.
7. Litter, plant roots, and the humus of forest soils aid porosity, reducing run-off and enhancing penetration.
8. Vegetation absorbs various minerals from the soil and upon decomposition, returns mineral elements to the soil.
9. Vegetation takes large amounts of water from the soil and transfers it to the atmosphere through transpiration.
10. Under certain conditions vegetation enables production of organic subsubstrate and organic soils, for example, in bogs, swamps, marshes and flats.

Other interactions occur. The examples are only illustrative, but they provide a basis for the generalization: 'if substantial changes are made to the vegetation of an area, there will very likely also be changes in the soil'. It will be recalled that this is what happens when tropical forests are removed for agricultural development. In a few years the soil becomes unproductive and must be rejuvenated by means of forest or bush fallow.

Over-all, the above statements are probably least appropriate for desert soils. Arid soils contain only about 0.2 to 0.5 per cent organic matter, and semi-arid soils only 1 to 2 per cent, so the usual role of vegetation on these soils is comparatively insignificant.

Man, through his various activities, may exert either a beneficial or a deteriorating effect on the soil. He may, through his actions, convert natural vegetation areas into useful agricultural land or he may accelerate erosion, through practices such as uncontrolled burning, overgrazing or cultivation of steep slopes without proper water management.

Land management: soil fertility

Earlier in this chapter comment was made upon the basic nutritional requirements of plants in relation to the soil, in so far as the inorganic elements are concerned in the healthy growth and development of the plant. In the maintenance of soil fertility, under continuous usage on an agricultural or small-holding basis, methods other than reliance upon chemical fertilizer application, are desirable. Thus, while not under-valuing the use of chemical fertilizers, rather encouraging their use with care and discretion, there are alternative ways of restoring nutrients to the soil and the more generally applicable

are green manuring, the use of legumes, crop rotation, animal manuring and the use of human wastes. Proficient soil-fertility maintenance would make use of a combination of these other ways, in addition to any application of chemical fertilizers, rotations and mized cropping.

When sufficient moisture is available in the soil both for the growth of plants and for reasonably rapid decomposition of ploughed-in plant matter, then green manuring can be of economic advantage. Further, provided the soil is in a reasonable physical condition, not too acid in reaction and in an area of good, seasonal rainfall, a green manuring will contribute towards the maintenance of fertility and conservation. Crops for the production of green manure must be grown frequently within the rotational system, if above-average crop yields, in general, are to be maintained.

It is unlikely that soil fertility can be maintained solely by green manuring and less frequent usage of this technique is recommended along with other measures, as is done by the peasant farmers in various parts of India and China. Green manuring is not popular with small-scale farmers in tropical areas, probably because their cultivable soils are of limited extent and it is thought to be wasteful of such soils in occupying the land with crops that are going to be ploughed into the soil, rather than harvested and eaten. Thus economic and labour problems, particularly in subsistence farming, also determine the negative attitude towards green manuring as an agricultural technique.

The faeces and urine of domestic animals are a good source of soil nutriment. In natural ecosystems, both the plant and the animal communities interact in respect of nutrient contributions to the soil; in ecosystems modified by man and particularly in land areas where soils are taken over for crop production and livestock farming, then animal excrement should be applied to the soil, to help maintain a good level of fertility. Manure, if processed prior to usage, should be covered i.e. not exposed to the hot sun, which destroys a great deal of the valuable combined nitrogen by heat and photoxidation. Table 37 illustrates the amount of manure available, assessed on the basis of 450 kg live weight, *per se*, for five types of farm animals. The table also shows the nutrient status of fresh manure, in each case.

Table 37. Relative amounts produced and nutrient status of animal manure[1]

Animal	Metric tons excreted per year	Percentage of						Dry matter kg per	
		Urine	Faeces	Water	Nitrogen	Phosphorus	Potassium	ton	year
Horse	8.1	20	80	78	0.71	0.105	0.514	218	1 782
Cow	12.2	30	70	86	0.49	0.061	0.394	137	1 701
Swine	13.8	40	60	87	0.38	0.148	0.348	129	1 790
Sheep	5.7	33	67	68	1.03	0.153	0.821	317	1 814
Poultry	3.9	—	—	55	1.00	0.349	0.332	446	1 742

1. Quantities of manure excreted by 450 kg live weight of different animals and the nutrient content of the manure (fresh manure without bedding).

Source: C. E. Millar. *Soil Fertility*. New York, Wiley, 1955, p. 332.

Well-rotted manures are best applied shortly before the planting of a crop and they never should be applied in the dry season, because their full nutrient potential will be much decreased by oxidation and leaching. Such manures should be used in association with carefully planned crop rotation. There is an implication, vis-à-vis animal manures and crop production, that good use of the land in farming means an integration of crop and animal husbandry.

Most animal manure sources are comparatively low in the element phosphorus. In many areas, low phosphorus becomes a limited factor in crop production, when animal manure is used by itself. Supplementary applications of phosphorus are recommended.

Human excrement

The contribution of human excrement to soil fertility has been recognized since ancient times throughout the civilized world. In many areas, human excrement continues to be returned directly to arable soils, though it is now known that certain parasites and pathogens such as *Entamoeba histolytica* and hook-worm (Nematoda) are transmitted through human faecal matter, causing the contamination of vegetable plants grown in the fertilized soil.

Modern methods of sewage treatment in high density urban areas produce a material called 'sewage sludge'. Table 38 provides illustrative data on the nutrient status of this material. The specific content of sludge varies considerably, based partly on the various methods of sewage treatment by which it is processed.

Table 38
Nutrient status of digested sewage sludge[1]

Element	Number of samples	Average	Range
Nitrogen (%)	26	2.07	1.30– 2.94
Phosphorus (%)	30	0.71	0.27– 1.92
Calcium (%)	30	2.49	0.68–10.00
Magnesium (%)	30	0.23	0.14– 0.45
Potassium		(none detected)	
Carbon-nitrogen ratio	16	13.2	10–22
pH	19	5.77	4.5 – 6.85

1. Unpublished data. Courtesy of H. A. Lunt.

Source: C. E. Millar. *Soil Fertility*. New York, Wiley, 1955, p. 325.

Some studies have shown crop damage due to the application of sewage sludge to acid soils. The addition of lime to such soils may be necessary for enhanced fertilizer effects to be realized.

Inorganic fertilizers

Commercially produced inorganic fertilizers are increasingly used in technological communities. The most common components are nitrogen (as ammonia, ammonium

salts, urea or nitrate), phosphorus (as rock phosphate, bones, phosphoric acid, calcium meta-phosphate, nitrogen phosphate or superphosphate) and potassium (as potash, potassium chloride, potassium sulphate or potassium nitrate). These are marketed as NP or NPK fertilizers and they have certain advantages for rural farming, since they are cheaper to pack, handle and transport, being concentrated granules. Storage quality is also ensured by virtue of the method of manufacture. There are also granulated compound fertilizers on the market but these should be used only after local soil nutrient requirements have been competently assessed by a soil scientist. The majority of tropical soils have very low phosphate levels and in most instances this element is not available for optimal plant growth. There are many examples of increased yields in response to the careful application of phosphatic fertilizers, particularly with yams, sweet potatoes, cassava, cotton, cereals and legumes. The commonest and most readily available phosphatic fertilizers used in general agriculture are the super-phosphates (with free, water-soluble phosphates). The double and triple super-phosphates with 42 to 49 per cent P_2O_5 are the most economical, although the initial cost seems to be generally higher than the ordinary super-phosphates. Rock phosphates, if available locally, are recommended; they release the phosphate after mixing with moist soil, with which they react, particularly in acid soils.

Alongside the phosphorus requirement of green plants is that of certain other elements, of which soils in Asia are generally deficient. One of these is sulphur and another is calcium. Even elements like magnesium may be quite deficient in soils of the humid tropics, but this does not have much adverse effect on annual crops. In tropical soils that are intensively cultivated, it is often advisable to determine the sulphur availability and to rectify any deficiency by the addition of the element as sulphate.

It must be remembered that our knowledge of the nutrient requirements of different crops grown on soils in Asia is still very limited, and the inconsiderate use of inorganic fertilizers, as a panacea for the increased productivity of the soils, is hazardous. Nutrient imbalance can be disastrous. An application of combined fertility-improvement practices, adapted to local requirements, is the safest solution to the problem at this stage.

Soil maintenance and problems in the tropics and subtropics

Much of the data on soils in temperate regions cannot be applied directly to tropical and subtropical soils, and comparatively little data is available about these soils. Furthermore, an international soil classification system has not become broadly utilized, so it is difficult to discuss soil maintenance problems in terms of major soil groups. Several major soil groups commonly recognized, however, are: desertic and sierozem soils (arid and semi-arid), lateritic soils (or latosolic or ferrallitic soils, also called oxisols), alluvial soils, volcanic soils. Alkalized and salinized (halomorphic) soils are often grouped with arid and semi-arid soils. A fifth and smaller group is composed of the dark grey and black soils, including the black cotton soils of India.

Since the desertic and sierozem soils are generally limited in productivity because of an inadequate moisture content, the use of fertilizers is generally not relevant unless irrigation is being used. Fallowing systems are sometimes used to preserve water for the following season's crops. Contour ridges with dams across the furrows every metre or so (called 'tieridging') are sometimes used, to prevent run-off and erosion and to preserve moisture. Irrigation is effective when the soil has high permeability, a deep water table, and exists on flat land. High permeability and the deep water table are important in allowing leaching and preventing alkalization and salinization. Even soils that have become saline can sometines be improved by the addition of fresh water, leaching and drainage. Alkalized soil, with its aqueous solution high in Na^+, can often be improved by the addition of gypsum ($CaSO_4 . 2H_2O$), resulting in the replacement of Na^+ on the adsorbing complex of the soil by Ca^{++}, thus improving soil texture.

In the semi-arid regions of Asia, drought-resistant crops such as new varieties of maize are receiving increased attention. The new maizes used in semi-arid regions, thus far, are effective because they are short-season varieties, and thus not truly drought-resistant. These often provide a crop during limited rainfall periods by avoiding the dry season when local slow-developing varieties may result in a complete crop failure. Research is continuing on the development of true drought-resistant varieties of maize.

Some shifting cultivation and fallowing is practised in semi-arid regions, but this must be done with due consideration since natural vegetation often causes a greater moisture depletion in soil than that caused by a short-season, cultivated crop. The nutrients found to be most likely to limit productivity in the semi-arid regions of Asia are nitrogen and phosphate. Careful choice of crop varieties, timely planting in terms of moisture and occasional judicious use of fertilizer are recommended.

Lateritic soils (latosols), so common in the wet, tropical regions, are old and deeply weathered soils, which are generally of low fertility. They lack weatherable minerals and have poor chemical properties, though their physical soil-properties are often quite good. They are porous, very resistant to erosion and neither shrink nor swell. Root growth and development in these soils is often limited. Detracting factors include their high aluminium content which is sometimes toxic to plants, low organic content (except soil newly reclaimed from tropical forests), low nitrogen, phosphate and potassium in many cases. It is hoped that the chemical requirements of these soils will be met in the future through a discrete use of fertilizers, but information is still inadequate since the fertilizer techniques of temperate climates cannot be directly applied. In West Malaysia, such soils have been improved for the cultivation of rubber by the use of fertilizers.

At the present time, the prevalent practice of shifting cultivation can still be considered as a solution to the problem. Generally, the accumulated evidence suggests that savannah regrowth is less effective than forest regrowth in restoring the nutrient status and in positive contributions to soil structure. Accordingly, forest fallows are recommended in shifting cultivation. In some cases, short-term fallows may provide adequate humus, but nutrients (especially phosphates) have then been added.

Alluvium (alluvial soils) is a very fertile soil and important in many parts of the world. If floods occur fairly frequently, the fertility is maintained by the new layer of silt brought in by the flood waters. Rice is the most important crop grown on such soils.

Volcanic soils are not always as fertile as is commonly believed. Such soils, when acidic in reaction, are comparatively infertile; conversely, volcanic soils which are basic in reaction are usually fertile.

A number of the dark grey and black soils (vertisols, grumusols) are characterized by extensive cracking and swelling. These are heavy textured, dark, clay soils which are less common than the above three groupings. Their swelling and cracking imposes an agricultural limitation since root systems may be damaged. Other community problems arising from this characteristic include cracked walls in buildings, leaning of poles, fences and trees, and cracking of irrigation and drainage canals. The dark grey and black soils often have a low permeability and are hard to cultivate because of hardness in the dry season and stickiness in the wet season. The impermeability also results in loss of much water due to non-absorption and to run-off. With carefully regulated irrigation, some crops such as cotton, tobacco, rice, wheat, sorghum, lucerne and sugar-cane are being grown on such soils. High fertility, uniformity and occurrence in flat regions are advantages.

Andosols are favoured for the growth of cinchona, coffee, irrigated rice, oil palm and tobacco in some areas of Asia.

Podsol soils are not very important agriculturally as they are highly acidic and coarse in texture. Coco-nut palms, citrus fruits, mangoes and cashew nuts have been successfully grown on these soils.

Grey hydromorphic soils are often used for growing rice in Asia.

Soil and human populations

We have already noted various examples of the interaction of human populations with the soil. These and additional examples are now to be discussed. The situations involving human populations generally pertain either to agriculture or to urban development. In many comparatively underdeveloped, rural situations, soils also influence human population density.

Agriculture

Sierozems and desert soils are often of limited agricultural value because of scarcity or sporadic nature of rainfall. Such soil may be useful for grazing, but it is readily overgrazed or overcultivated. These conditions lead to soil damage by wind erosion and decreased productivity. Irrigation may dramatically increase production, but salinization and alkalization may be problems unless soil drainage is good.

The conversion of tropical forests, woodlands and savannahs to areas of agriculture

often results in a rapid drop in productivity. Shifting cultivation and crop rotation have been the responses made by human populations to this problem. Each of these will be considered respectively.

Shifting cultivation is still very widespread in practice in Asia, as well as in tropical America, Africa and Oceania. Land is cleared of natural vegetation, whether it be forest or savannah. The debris is burned and the plant ash is worked into the soil. Crop plants, appropriate to the traditions and customs of the community are grown and cropping continued for various periods (between two and four years). The land is then abandoned and may lie fallow, with a regenerated natural flora developing over the years. New areas are similarly cleared and prepared for cultivation. In shifting cultivation patterns no fertilizers or manures are put on the soil, and even if grazing herds are kept by the farmer, no use is made of the animal excrement. Subsistence farming would be satisfactory for small populations, as indeed it was done over previous centuries; it did little harm to the soil, because the fallow periods were long enough to bring about a restoration of soil fertility. In drier woodland and grassland zones, however, the fallow periods do not restore fertility to the same suitable degree as in forest zones and bush fires, started in such open areas by man, spread so easily and do a great deal of damage to the fringing vegetation and, incidentally to the wild fauna, depriving them of food material and habitat. With increasing human populations, the intensive shifting cultivation now practised for the growth of cash crops as well as subsistence farming, leads to a serious decline in soil fertility. The principal vegetational areas in which shifting cultivation is practised are in woodlands and savannahs, in semi-deciduous forests, in lowland evergreen and semi-evergreen forests and in the mountain forests.

Those communities which practise shifting cultivation under different topographic and natural vegetational conditions are very good examples of the interaction of human populations with the environment. Village life and customs are very dependent upon the response of the soil to their demands, hence human settlement patterns become indicators of the status of soil-man relationships.

The population of South-East Asia is unevenly distributed. In some areas the densities are over 575 per square kilometre while other areas are sparsely populated or totally uninhabited. The pattern of distribution is governed largely by the potential for productive agriculture in the lowlands and deltas. Here the physical environment is conducive to agriculture as illustrated by the valleys of Irrawaddy, Chao Phraya, Mekong, Salween, Sittang and Red Rivers, the fertile coastal plains and in areas of fertile volcanic ash, especially in Java. Sedentary agriculture like the cultivation of wet padi, supports a high population density as well.

By contrast, shifting agricultural practices do not support large populations and have low-population densities. Inaccessibility, rugged relief and poor soils may also account for low-population densities as in the interiors of Borneo and Burma.

Decreasing land productivity is reflected in the migration of whole communities and the opening up of new areas. It is obvious that a heavy depletion of certain mineral elements will occur where this type of cultivation is based upon a single species of crop plant, which forms the staple food of a community, e.g. bananas, plantains,

cassava, millet. Patterns of land ownership and usage are also influenced by the *modus operandi* and *modus vivendi* in areas subject to shifting cultivation. Thus there is a need to modify or replace shifting cultivation patterns with systems that will maintain, improve and increase soil fertility and enhance productivity. Crop rotation and mixed cropping systems can be of use in replacing shifting cultivation. Crop rotation or bush fallow deliberately encourages a natural vegetation re-growth to restore soil fertility after short periods of cultivation. We must bear in mind that the succession of crops on a villager's farm is not with any variety, rather the cultivation, each season, of a staple food plant. So, suggesting a change from shifting cultivation to rotation not only demands a change in attitudes towards farming but also towards diet.

Erosion

The immediate causes of erosion are usually water, wind, ocean waves or glaciers. However, soil modifications initiated by man are often major ultimate causes of increased rates of erosion by water and wind. In nature, vegetation inhibits erosion through such processes as the following: roots held soil together; leaves intercept rain, breaking impact of water on soil; plants shield the surface of the soil from wind and solar energy; plants provide humus. The implications of the above actions are numerous. The surface porosity of the soil is maintained by reduced impact of water on soil and by humus. Similarly, percolation is aided and run-off limited. The rate of drying of the soil is inhibited as is also the movement of soil particles.

When soil is mismanaged by man with resultant increase in erosion, it is usually because of overcropping, inappropriate ploughing, overgrazing or deforestation. Overcropping reduces humus with resultant loss of soil cohesiveness and permeability. Inappropriate ploughing exposes soil surfaces to direct action of the wind and water, resulting in greater likelihood of particle movement. On steep slopes sheet erosion followed by gully formation are results of action by water. In arid regions, dust storms and sand storms result in removal of rich topsoil. Overgrazing and deforestation directly remove vegetation with resultant loss of the protective mechanisms mentioned above.

Examples of erosion due to human activity are available throughout the world and the following are only illustrative of the Asia region.

One of the agricultural systems still practised in South-East Asia is that of shifting cultivation. First, a clearing is made in the forest by cutting and burning some trees. Then crops are grown on the cleared land. After two or three harvests, when the soil fertility is depleted, the land is abandoned for another new area. Soil erosion has often been a problem on exposed agricultural surfaces, especially on steep slopes.

In India, deforestation and overgrazing have been causes of erosion. In the Punjab Siwaliks, for example, deforestation has been followed by the grazing of hillsides. This combination has resulted in removal of soil-protecting vegetation. In China, the erosion and periodic flooding of the Yellow River are noteworthy. In Indonesia, shifting cultivation, unrestricted grazing and burning of forests and grasslands are threats to soil stability.

Irrigation

Irrigation has been used and developed by a number of civilizations, with great efficiency over the last 6,000 years. Today, vast areas of the best agricultural land is irrigated, including all of the cultivated land in Egypt. Irrigation, as an agricultural technique for increasing soil productivity, has proven successful in areas of high population living along river systems and around river deltas.

Irrigation is the controlled usage of water on crops and soils. It requires careful planning in order to ensure the most economical and productive use of the water without causing erosion, increased salinity or over-saturation of the soil. The earliest types of irrigation projects used the seasonal field waters of the land along riversides by impoundment; such is possible in areas where the land is flat. Similarly, water can be drawn from the river, either mechanically or by gravitation, and run into impoundments or channels, for further distribution. The ridging of soils, either along contours or slightly obliquely across contours and ridges can be used to control and make improved usage of water run into the system. In any event, there is always the possibility of soil erosion due to the irrigation becoming uncontrolled and systems must be devised to ensure adequate soil conservation while benefiting from the enhanced water supply.

Alongside modern technological developments such as man-made lakes, hydro-electric schemes and land reclamation schemes, much use can be made of the impounded or restricted waters in the development of controlled irrigation. Indeed, a feature of some of these large-scale projects is the improvement of land use, soil fertility and productivity in areas adjacent to dams and other impoundments.

Underground irrigation has been achieved with success in certain countries such as the United States and in South Africa, using subterranean perforated, polyvinyl plastic tubing, through which the water is pumped under pressure and sprays out into the soil in all directions. This system has the advantage of reducing water loss by evaporation from the soil surface which does occur in any overground application. In hot, dry climates, therefore, this technique would be recommended. It can be costly in the initial outlay, but with careful planning it could pay for itself in increased crop production and market profits.

Irrigation by sprinkling, using rotating sprinkler-heads adjusted to height and distance of spray, is used with success in the production of fruit, vegetables, tobacco and sugar. The water supply originates either in an underground source (aquifer) or can be pumped mechanically from nearby rivers, streams, ponds or lakes. Sprinkler methods can also be used to apply soluble fertilizers to crop plants as foliar nutrients, these being allowed to enter the water supply from specially constructed feeder-reservoirs; this use ensures the alleviation of mineral deficiency symptoms in both the soil and the plants growing in it.

Fig. 75 illustrates some methods of irrigation in present-day use.

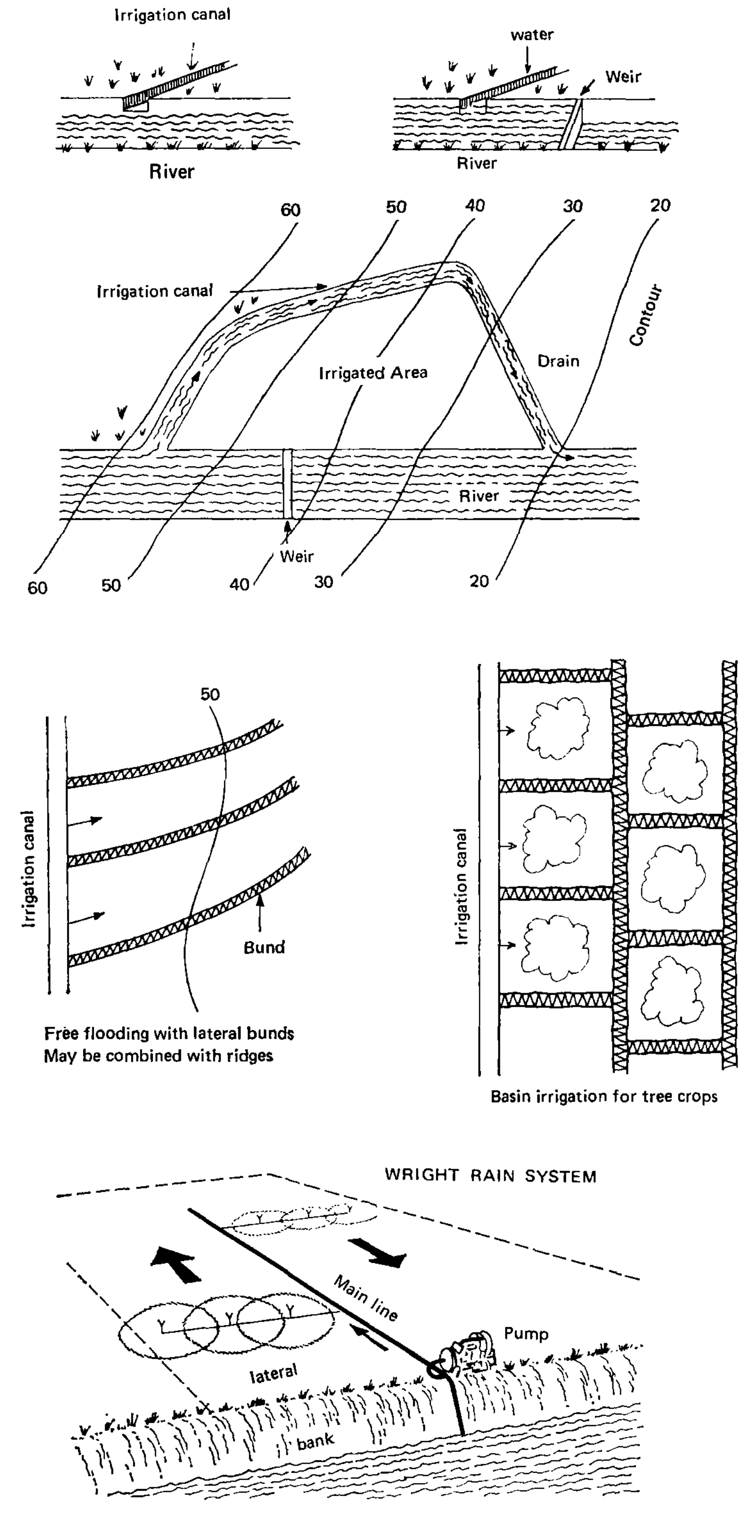


Fig. 75
Some methods of irrigation in present-day use
Source: G. Wrigley. *Tropical Agriculture*. London, Faber, 1969, p. 145–8.

Rural development

Rural development is, in effect, the improvement of the standards of living of villages and hamlet communities, particularly through the evolution of agriculture and land-use reforms.

In any programme involving the re-location of settlers, four things must be borne in mind. First, it is necessary to establish some law and order in the rural areas to counteract robbery and extortion. Second, there is the need to ensure an adequate supply of water for domestic and farm use because in this way the rural population does not have to depend on the inadequate water supplies from streams, lakes and wells. Third, it is necessary to eradicate diseases prevalent in rural communities, such as malaria, sleeping sickness, bilharzia, thus to ensure better human health standards. Fourth, there is a need to institute land-reform legislation and agricultural improvement.

In re-locating rural settlements it is necessary to reduce the unnecessary time spent in communicating between fields and hamlets. Settlements must be constructed in close proximity to farm-lands. From the agricultural point of view it is important to irrigate large areas of land and to introduce more intensive methods of cultivation. Rearing of livestock and poultry should be encouraged. Further it is necessary to provide good roads, bridges and processing centres if the rural people are to be brought into the national schemes of agricultural productivity and trade.

The above problems and their solutions seem to work for most countries but in a small country like Singapore there is another important problem and that is the scarcity of land in resettlement areas. The Government of Singapore has come up with a partial solution that is to use reclaimed swamp-land for the growth of vegetables. At the same time the government is looking into the agricultural potential of off-shore islands like Pulau Ubin and Pulau Tekong which jointly offer something like 2,800 hectares of land.

Urban development

Soil permeability and stability are important in urban development for several reasons. When soil is relatively impermeable to water, there is greater likelihood of flooding in lowlands. Thus, settlements should be restricted to higher elevations than is necessary where soil is highly permeable. Home septic tank filter fields, when present, will not function effectively in impermeable soils. On the other hand, highly permeable soils or fissured limestone may not be desirable either. Under these conditions, there is danger of rapid entry of sewage into groundwater supplies. Thus, as human populations and urbanization increase, more advanced methods of sewage disposal are generally required.

Soil stability must be considered in the placement of buildings and roadways. On slopes, soils with poor stability may slide when saturated, resulting in damage to the homes and possible loss of lives. Slippage or settling of soil may also occur in lowland areas where soils are unable to support the weight of roadway usage or buildings.

Land classification and management

Land classification and use

Classification of land may be based on any of numerous criteria, which are derived by the disciplines of economics, agriculture, engineering and town and country planning. The economist is interested in the over-all productive potential of the land while the agriculturist is also concerned with productivity but from the perspective of crop production and animal husbandry. The engineer, on the other hand, is more interested

Table 39. Outline of land-capability classification

Major land-use suitability (Broad grouping of limitations)	*Land-capability class* (Degree of limitations)	*Land-capability subclass* (Kind of limitations. Grouping of land-capability units). Example of possible subclasses in class III:	*Land-capability unit* (Distinctive physical characteristics. Land-management groups based on permanent physical factors). Example:
	I Few limitations. Wide latitude for each use. Very good land from every standpoint.		
	II Moderate limitations or risks of damage. Good land from all around standpoint.		
Suited for cultivation	III Severe limitations or risks of damage. Regular cultivation possible if limitations are observed.	Limited by hazard of water erosion; moderately sloping land.	13–C–2 Moderately sloping, slightly acid soils on limestone.
			9–C–2 Moderately sloping, highly acid soils on sand stone or shale.
		Limited by excess water; needs drainage for cultivation.	
		Limited by low moisture capacity; sand land.	
		Limited by tight very slowly permeable subsoils, claypan land.	
	IV Very severe limitations. Suited for occasional cultivation or for some kind of limited cultivation.		

Major land-use suitability (Broad grouping of limitations)	*Land-capability class* (Degree of limitations)	*Land-capability subclass* (Kind of limitations. Grouping of land-capability units). Example of possible subclasses in class III:	*Land-capability unit* (Distinctive physical characteristics. Land-management groups based on permanent physical factors). Example:
Not suited for cultivation; suited for permanent vegetation.	V Not suited for cultivation because of wetness, stones, overflows, etc. Few land limitations affect grazing or forestry use.	Grouping of range, pasture or forest sites.	Land-management groups based on permanent physical factors, such as range sites or forest sites.
	VI Too steep, stony, arid, wet, etc., for cultivation. Moderate limitations for grazing or forestry.		
	VII Very steep, rough, arid, wet, etc. Severe limitations for grazing or forestry.		
	VIII Extremely rough, arid, swampy, etc. Not suited for cultivation, grazing or forestry. Suited for wildlife, watersheds or recreation.		

Source: United States of America. U.S. Government Publications. *A Water Policy for the American People.* Report of the President's Water Resources Policy Commission, Vol. 1, 1950.

in the suitability of the land for building dams and highways while the city planner may be concerned about the degree of stability of the soil and substrata to support building structures.

Soil scientists have now developed classification systems based on the capacity of land to support and maintain permanent agricultural exploitation. Such a system of classification is called a land capability classification. Table 39 gives such a classification and Table 40 shows land-use in Asia.

An initial activity required in any land classification system is the construction of a suitable land map, using serial photographs of the region. Pertinent information such as kind of soil, percentage of slope, degree of erosion, flooding hazards, condition of wetness, alkaline or salinity problems and gross texture are recorded as accurately as possible on the map. In addition the present use of the land (cultivation, pasture, forests, etc.) is also indicated.

A land classification system to be used in developing countries generally categorizes land suitable for cultivation from land not suitable for cultivation. In the system discussed here, the land that is suitable for cultivation is further classified into four classes (I, II, III & IV), as is the land that is not suitable for cultivation (classes V, VI, VII, VIII). These eight classes of land exist in almost all countries, thus the descriptions which follow are widely applicable by farmers, teachers and soil scientists.

Table 40. Land use

Continent or country	Total area	Agricultural area: Arable land	Agricultural area: Land under permanent crop	Agricultural area: Permanent meadows and pastures	Forest land	Other land
World	13 391 000	1 432 000		3 059 000	4 028 000	4 872 000
Africa	3 030 000	205 000		844 000	639 000	1 342 000
N. America	1 968 000	220 000		280 000	739 000	729 000
L. America	2 056 000	119 000		505 000	984 000	448 000
Asia	2 753 100	461 300		499 000	534 600	1 258 000
Europe	493 000	147 000		93 000	140 000	113 000
Oceania	851 000	47 000		464 000	81 000	259 000
U.S.S.R.	2 240 220	227 800	5 009	374 000	910 009	423 402
Asian countries						
Afghanistan	64 750	7 844	136	6 020	2 000	48 750
Burma	67 803	18 488	453	375	45 274	3 213
Ceylon	6 561	895	1 084	439	2 899	1 244
China	956 100	110 300		177 000	76 600	592 000
India	326 809	160 540	4 070	13 880	62 320	85 999
Indonesia	190 435	18 000		—	121 800	—
Iran	164 800	16 060	500	11 000	18 000	119 240
Iraq	43 492	10 000	163	63	1 851	31 415
Israel	2 070	337	86	818	109	720
Japan	36 988	4 910	600	948	25 558	4 972
Jordan	9 774	1 132	168	100	125	8 249
Kuwait	1 600	0.5	—	134	2	1 464
Lebanon	1 040	240	76	10	95	619
Malaysia (W)	13 131	600	2 257	—	8 037	—
Philippines	30 000	6 574	2 403	1 423	15 899	3 701
Saudi Arabia	214 969	765	44	85 000	1 680	127 480
Syrian Arab Rep.	18 518	5 641	258	5 434	440	42
Taiwan	3 596	322	546	2	2 319	408
Thailand	51 400	9 746	1 669	—	27 354	—
Turkey	79 058	24 793	2 585	26 135	18 274	6 272

1. All data of table are in 1,000 hectares. *Source: FAO Production Year Book*, 1971.

Class I. Land indicated as class I has little or no limitations in use. It often consists of a field that is almost level and has deep and fertile soil. The soil is generally free of gravel and stones and is not subject to wind erosion or water erosion. A class I soil has sufficient rainfall for the growth of nearly all crops. It never gets flooded and has no drainage or alkaline problems. A major change in any one of the above conditions automatically shifts the land from class I to class II. Class I land requires good soil management practices such as crop rotation and the addition of commercial fertilizers. With good soil management such land can be farmed indefinitely.

Class II. Class II land is characterized by soil having a depth of about 1.0 m and a 5 per cent slope. The soil is fertile, usually of the silty or loamy type. To obtain

yields comparable to those from class I soils, the land in class II must be managed in such a way as to prevent erosion and soil damage. For example, practices should be designed to include contour cultivation and cross-slope seeding of grains and legumes.

Class III. Class III land is characterized by light, sandy soils on slopes steeper than those belonging to either class I or II. Such land has hazards or limitations that may be drastic to crop growth if not properly managed. At some time in the past, these lands may have had very much rich soil that has been allowed to erode away. Such land requires manuring, commercial fertilizers and a combination of strip cropping, terracing or other practices to maintain good yields and at the same time to prevent erosion. Class III land may also include shallow soils on nearly level slopes or poorly drained lands which require careful management.

Class IV. This class includes most of the shallow sandy or heavy clay soils with steep slopes. Eroded soils that can be successfully cultivated belong to this class. Unlike the land in classes I to III, the land in class IV must be used for close growing crops such as pasture or hay most of the time. This class includes areas of land with high wind velocity, low moisture zones and with a short growing season.

Class V. Class V land is nearly level land with adequate moisture and soil deep enough to grow grass but too stony to support other vegetation. This land, as well as the land of classes VI and VII may be used for grazing domestic livestock or wild game or for growing trees. It is likely to suffer from erosion unless there is good forestry management.

Class VI. This class includes good grassland and forest land with slopes as high as 40 per cent. This land cannot be cultivated as it may be too stony, too steep, too sandy or too gravelly. The climate may be very dry so that the land is too arid for cultivation. However, where there is a deep soil which is well drained, orchards and vineyards and similar plant production may be developed. In such cases, perennial ground cover crops need to be planted to prevent erosion. Many areas of class VI land have been used as pasture for livestock.

Class VII. This land is very steep, with slopes varying from 40 to 65 per cent. The slopes are badly eroded and possess only a thin layer of soil. Such land may still be used for grazing or for timber production provided soil management is improved.

Class VIII. This class includes all land having slopes of 65 per cent and over plus land which cannot be placed in any of the other seven classes. This land is very badly eroded, rocky, steep, wet or arid and be used for wildlife cover. This class also includes swamp-land that cannot be drained.

In India about 30 per cent of the land is uncultivable due to topography, relief and the inadequacy of rainfall. Of the remaining 70 per cent, a large portion is required for the various projects mentioned above. Thus the land utilized to satisfy human needs amounts to only about 40 per cent of the total surface area of India. This was estimated to be approximately 130 million hectares in 1960–61.

Similarly, only about 30, 27 and 17 per cent of the total land surface in West Malaysia, the Philippines and Japan respectively are cultivable. Singapore has a limited area of 545 square kilometres and does not have much land for agricultural purposes,

but Burma has almost 9 million hectares of cultivable land. Java alone has about 65 per cent of the total cultivable land of Indonesia.

Table 41 shows important cash crops grown in different countries of Asia.

Table 41
Different cash crops produced by Asian countries

Crops	Countries
Rice	Burma, Thailand, China and India
Wheat	India and Pakistan
Natural rubber	Malaysia, Indonesia, Thailand, Sri Lanka, Democratic Republic of Vietnam, Republic of Vietnam and India
Tea	India, Sri Lanka, Indonesia, Pakistan, China and Malaysia
Coco-nut	Philippines, Indonesia, Sri Lanka, India, Democratic Republic of Vietnam, Republic of Vietnam, Malaysia and Thailand
Oil palm	Malaysia and Indonesia
Spices	Indonesia, Malaysia and India
Sugar	India, Philippines, China, Indonesia and Pakistan
Coffee	Indonesia and India

The cultivation of cash crops is done either in estates or in small-holdings. Higher yields and quality products are obtained from crops grown in estates while those grown in small-holdings usually produce lower yields and poorer quality products.

It is interesting to note that in 1965 there were 1.9 million hectares under rubber crop in Malaysia of which 39 per cent were in the hands of estates and the remainder in the hands of small-holders. The proportion of production was 52 and 48 per cent respectively.

To grow rubber successfully requires moderately fertile soils, preferably crumbly, deep, well oxidized and acid in reaction, with the water table being deeper than 1 metre. Rubber also grows well on reclaimed swampy soils.

Rice, an important staple food of most Asians, occupies an area of over 4 million hectares in Indonesia, 5.26 million in Burma, 7 million in Thailand and 33–35 million in India. Rice requires soils having a high clay content. These soils must be impervious as flooding of the paddy fields is important during the growing season. It is interesting to note that India is able to have three croppings in a year on the same soil.

Wheat is the second most important cereal in India next to rice. It is grown on about 12 to 14 million hectares of silt and clay loams although medium to heavy textured fertile and well-drained soils are equally good for this cereal.

Palm oil, of which Malaysia is the world's largest producer, is cultivated on an average of 105,300 hectares. This crop does well on alluvial loamy soils with good drainage.

Coco-nuts, of which the Philippines is the world's largest producer, is cultivated on over 1.6 million hectares of land. These palms flourish on sandy coastal soils though they also grow on hills up to an elevation of 486 m.

Sugar-cane, only second in importance to coco-nuts in the Philippines, which is the world's largest producer. It grows best in soils of sandy loam which do not at any period become flooded or swampy.

Manila hemp, an important commodity of the Philippines, grows well on rich well-drained soils like the fertile volcanic soils.

Livestock

India has the world's largest bovine population with approximately 150 million cattle, 43 million buffaloes, 40 million sheep, 47 million goats and 1.5 million horses, which prove a strain on the fodder resources of the country.

In the Philippines, the rearing of livestock such as cattle and pigs as well as poultry farming is done either on an industrial basis or on a rural scale. The number of cattle stands at 1.5 million, the number of buffaloes at 3.5 million and the number of pigs at 7 million. The poultry number is around 1.9 million, and about 1.72 million hectares of grassland is used as ranches for the cattle.

On the other hand, Thailand padi farmers depend on their 6.8 million water buffaloes for ploughing their fields. There are in addition about 5.2 million bullocks and 12 million cattle which graze on the open fields. Pigs are also reared in Thailand besides 11,000 to 12,000 elephants used as loggers in the timber industry.

Timber

The production of timber is an important export industry in West Malaysia, the Philippines and Thailand. The timber is often sold in one of two forms—round logs and sawn timber.

About 70 per cent of the land surface of West Malaysia is covered with forests. These forests are of three main types, namely: the lowland rain forests, up to an elevation of about 620 m; the freshwater swamp forests, covering about 5,180 square km; the mangrove swamp forests, covering about 1.425 square km of tidal land found mainly in the west coast.

About half of the total land surface is covered with primary and timber forests. Among the forest products, *meranti* (light and medium hardwood) is the most important timber produced. These trees grow on lateritic soils to an elevation of 775 to 875 m.

Land management

Two major aspects of land management are the control or erosion (movement of weathered rock and soil by wind, water and ice) and maintaining the chemical and physical structure of the soil.

Erosion control. Severe erosion usually also results in the deterioration of the chemical and physical composition of the soil, further emphasizing the importance of this aspect

of human ecosystem management. An additional aspect of land management is the regulation of uses to which land is subjected.

Geological erosion occurs naturally at a slow rate creating such features as valleys, plains and deltas. Factors which influence the rate of erosion include the type of soil, vegetative cover, wind velocity, topography and amount and intensity of precipitation. Concern here is focused on control of accelerated erosion as these factors are modified by man and as the soil is mismanaged. Some common examples of mismanagement are overcropping, overgrazing, deforestation and cultivation of unsuitable soil. The major methods currently used for controlling accelerated erosion are contour tillage, terracing, gully reclamation, strip cultivation and wind-breaks (Figs. 76 and 77).

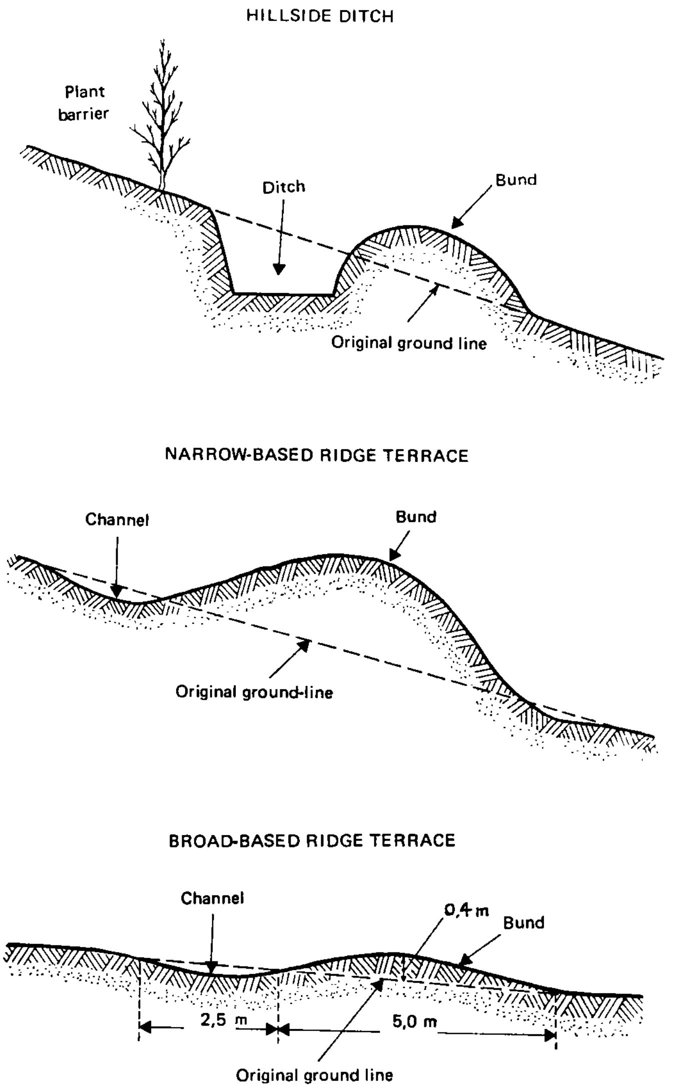


Fig. 76
Diagrams illustrating types of soil conservation procedures
Source: C. C. Webster; P. N. Wilson. *Agriculture in the Tropics*. London, Longmans, 1966, p. 110–12.

Contour tillage. If the slope is steep, the rapid downward rush of water results in soil movement and gully formation. With cultivation, the erosion process is often accelerated because, for part of the year, the soil retention by plants roots is reduced, and impact of raindrops is not broken by the vegetation or litter.

Contour tillage involves cultivating, seeding, and harvesting across a slope rather than above and below it. Furrows and ridges thus retain water, aiding soil penetration. Studies have shown reduced run-off on contour plots compared with non-contoured plots. In addition, the ridges serve as retaining walls limiting the downward movement of soil. Disadvantages of contouring include danger of water breakover in high rainfall intensity resulting in gully formation and reduced effectiveness when the soil surface

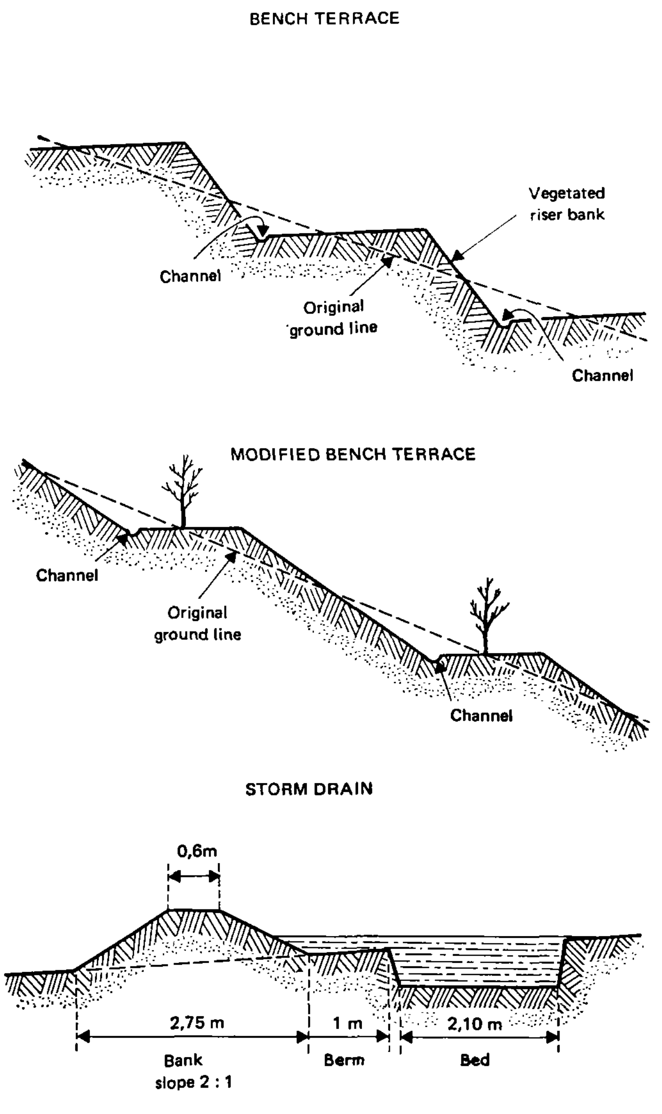


Fig. 77
Diagrams illustrating types of soil conservation procedures
Source: C. C. Webster; P. N. Wilson. *Agriculture in the Tropics*. London, Longmans, 1966, p. 110–12.

seals and water penetration is decreased. When contour tillage is practised on gently sloping ground and in low rainfall areas it is fairly effective in conserving soil and water.

Contour cultivation experiments done in certain states of India have given about 10 to 15 per cent improvement in yields, while also saving power in tillage operations. Contour trenching is often practised in grassland management.

Terracing. This term is applied to two quite different processes. The Incas of Peru and the Chinese have long developed and used flat, step-like shelves cut into hillsides and mountain slopes. Such flat shelves are used for dry farming in some regions of the world and for flooded rice cultivation in other regions. In the latter case, the course of water down the slope from one shelf to the next is regulated in channels for this purpose.

A second type of terrace has become widely used in modern agriculture. Such a terrace is an embankment or large ridge of earth, also called a 'bund'. It is constructed across the slope in such a way that it controls water run-off and soil erosion. Spacing between ridges must be close enough to check movement of water before it reaches high enough velocity to loosen and move soil (around 1 m per second). Other specifications which must be included in terrace design include terrace gradient (slope for drainage, if desired, but without erosion), terrace length in relation to field size and shape and desired terrace length in relation to field size and shape and desired terrace outlets, and terrace cross-section in relation to terrace water capacity and tillage convenience. Such terraces are effective on lands having slopes of up to 10 per cent. Level terraces (i.e. with zero gradient) do not require outlets, but if drainage is required, outlet location and design must be planned to minimize erosion. Natural drains, road ditches, tiled or otherwise-protected drains, and permanent pastures are among suitable alternatives.

In India, studies have shown that yields increased by 20 to 30 per cent through use of bunds. They also indicate decreased run-off on terraced plots against non-terraced plots.

Gully reclamation. If the land is badly eroded and gullies have formed, several procedures may be required for reclamation. A first step is to check the rapid flow of water down the gully. This is often done with small dams spaced appropriately along the gully to impede water flow (perhaps at intervals of 6–8 m, depending on slope). Dams may be made with brush, straw, stakes, logs, sods, wire netting or earth. The lower portion may be reinforced with rocks if a fairly substantial dam is required. As these dams gradually fill with sediment, the gully becomes filled in. An alternative method is to divert run-off from above the gully.

A second step is to hold the soil along the slopes of the gully. Straw may be spread and grasses or vines planted. Where soil types permit, litter and humus then begin to develop.

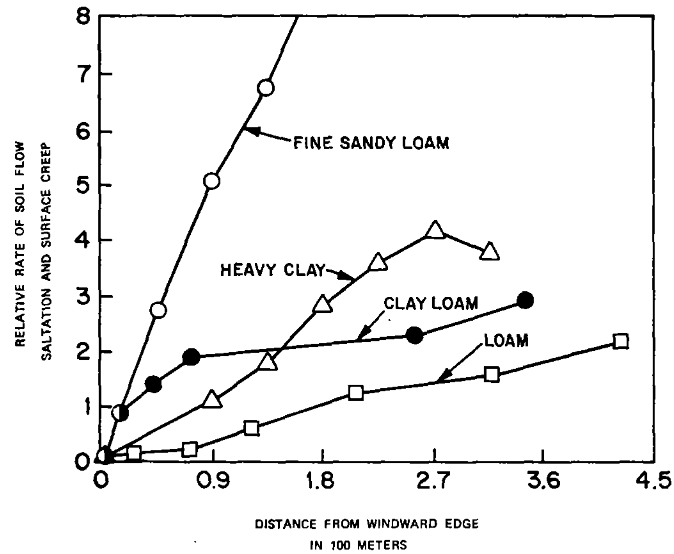
Strip cultivation. This technique is used as a means of controlling both wind and water erosion and as a method for effecting crop rotation systems. If there is a considerable slope to the land, strip cultivation may be combined with contour farming. Contour strips are then often arranged alternately or in a repeating sequence for crop rotation, including fallow strips.

Wind erosion is reduced since no large areas of soil are left bare at one time and since strips of varying vegetation types tend to serve as windbreaks for the areas of bare soil. If strips are straight, they should be across, rather than parallel with, the prevailing wind direction during the season of most severe wind erosion. Since soil vulnerability to wind erosion varies according to soil type, optimum strip-width will depend somewhat on soil type. For example, erosion of the heavy clay soil in 270 m strips would be about four times as great as in 90 m strips. When strip cultivation is used in combination with terracing, the strip width usually is made to correspond with the terrace interval.

Fig. 78 presents the relationship between length of eroding surface and soil movement for several soil types. These exact data would not necessarily apply to all soils of similar texture. However, with such data, decisions on strip width can be made.

Fig. 78
Soil erosion in relation to eroding surface length

Source: Richard K. Frevert; Glenn O. Schwab; Talcott W. Edminster; Kenneth K. Barnes. *Soil and Water Conservation Engineering*. New York, Wiley, 1955, p. 135.



In India, row crops like cotton and sorghum which permit erosion are grown in alternate strips with close-growing crops like ground-nut, other legumes, and grasses which limit erosion. Any run-off water or eroded soil that occurs in the soil-exposing strip is checked by the next soil-protecting strip. In other cases strips of anti-erosion crops like kidney beans and ground-nut 3.7 to 7.3 m wide often alternate with strips 22 m wide made up of erosion-permitting crops like sorghum.

Shelter belts. A shelter belt or windbreak is usually made up of one or more rows of trees placed strategically to reduce the wind velocity at the surface of the soil. Fig. 79 shows windward and leeward effects on wind velocity of a shelter belt at a height of 10.5 m. In general, the wind velocity is reduced to some degree on the windward side of the shelter belt for a distance of about eight times the shelter belt height, and on the leeward side, about twenty-four times its height. Further, a properly designed shelter belt of suitable dimensions may reduce a wind velocity of about 33.6 km per hour to about 12.8 km per hour to leeward.

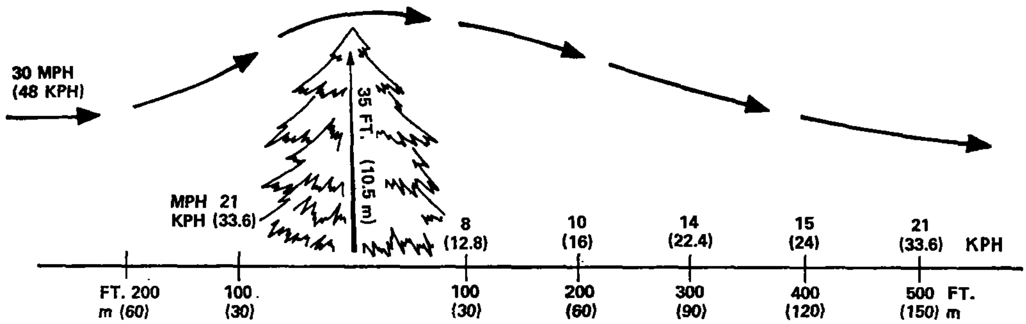


Fig. 79
Shelter belt and wind velocity
Source: Oliver S. Owen. *Natural Resources Conservation*. New York, MacMillan Co., 1971, p. 95.

Additional benefits from windbreaks include reduced evaporation (and therefore increased soil moisture retention), reduced firing or burning of crops from hot winds, and better fruiting in orchards. They also help to moderate local temperatures.

Tillage practices. Two significant tillage practices in the control of wind erosion are the production of a rough, cloddy soil surface and the retention or application of plant

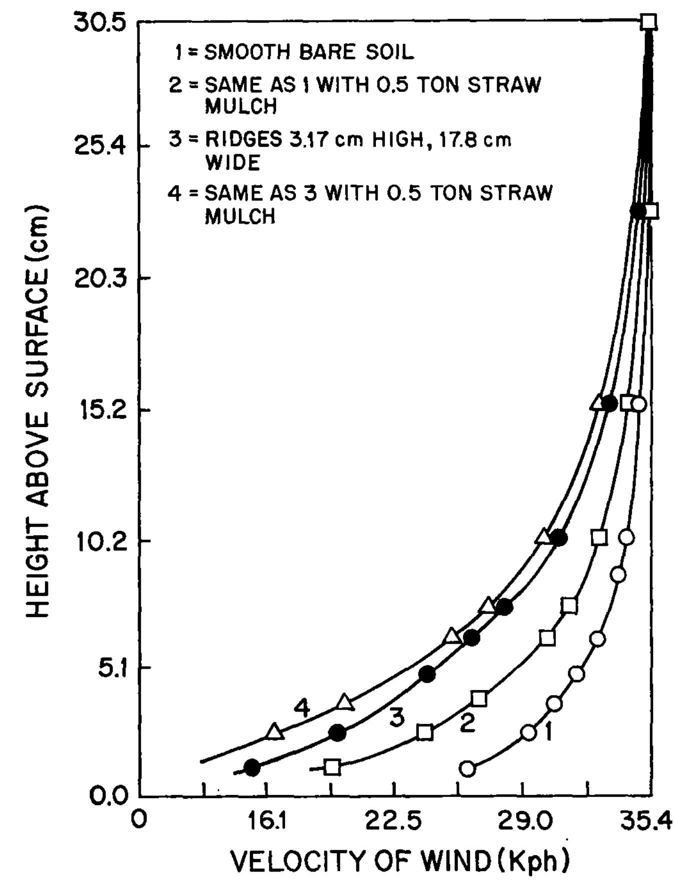


Fig. 80(a)
Tillage practices, wind velocity and soil erosion
Source: Richard K. Frevert; Glenn O. Schwab; Talcott W. Edminster; Kenneth K. Barnes. *Soil and Water Conservation Engineering*. New York, Wiley, 1955, p. 137.

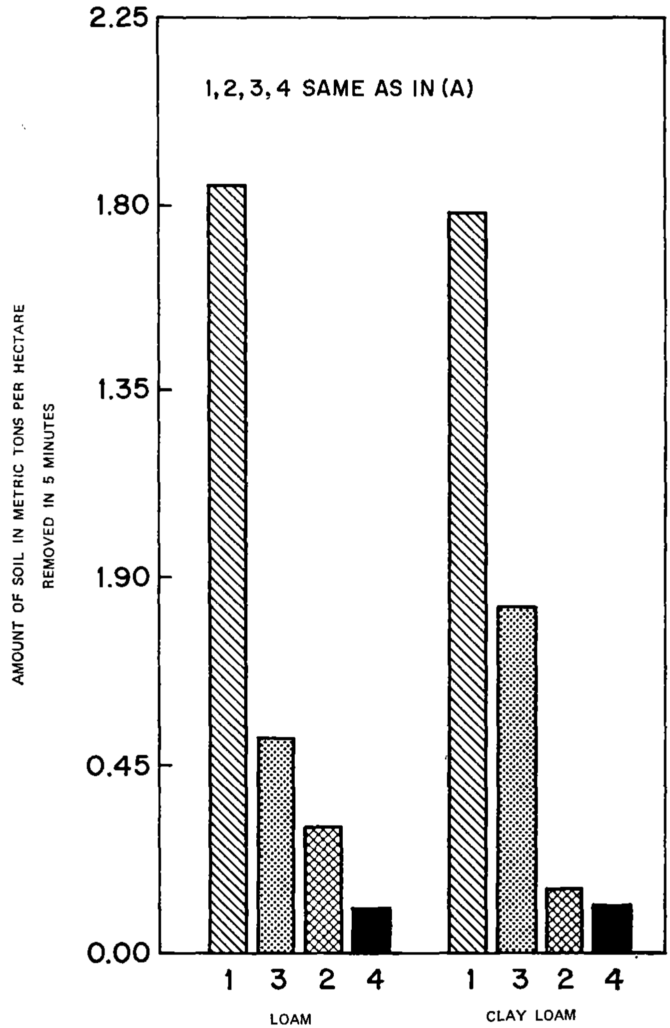


Fig. 80(b)
Tillage practice, wind velocity and soil erosion

Source: Richard K. Frevert; Glenn O. Schwab; Talcott W. Edminster; Kenneth K. Barnes. *Soil and Water Conservation Engineering*. New York, Wiley, 1955, p. 137.

remains or crop residues. The rough surface and the plant residues reduce wind velocity at the soil surface, and plant residues trap eroding soil. Fig. 80(a) (b) provides illustrative data on influences of these practices on wind velocity and soil erosion.

To illustrate, the data indicate that when ridges 3.17 cm high and 17.8 cm wide are used, wind velocity at about 7.6 cm above the soil surface is about 28 kmph, compared with about 34 kmph wind velocity over smooth bare soil. Where smooth, bare loam soil has about 1.8 metric tons per hectare removed by erosion in five minutes, loam soil tilled as described above results in about 0.45 metric tons per hectare removal.

Plant residues (leaves, straw, hay, sawdust) also break the fall of raindrops, impede surface flow of water, and promote infiltration of water into the soil. In warm and humid regions where decomposition is very rapid, it is often desirable to retain plant material in the living condition as long as possible so that it can serve as much for the new crop at planting time.

8 The biotic environment: flora and fauna

Introduction

The role of biotic factors in the operation of an ecosystem has been treated in the 'Introduction' and in Chapter 7 on 'soils'. Further discussions on man's relations with organisms are also included in Chapter 10 on 'food and nutrition', Chapter 11 on 'diseases and environment' and in Chapter 18 on 'ecosystem management and control of environmental quality'. In this chapter, the emphasis will be mainly on cultivation of plants, domestication of animals and biogeographical regions.

Man in the food web

Interactions between organisms in terms of food have been explained in the main 'Introduction'. Accordingly man is a consumer animal. Until he learnt to use fire he was mainly a consumer of second or third order, living primarily on animal food such as fish and insects. At times, however, he was a first-order consumer when he gathered and collected plant products (especially seeds and fruits) for food. Progressively, with the use of fire, man was able to incorporate many a species otherwise inedible into his regular food. As he left behind his apish ways he learnt to modify characteristics of organisms to his benefit through selection and (more recently) experimental breeding. Thus man has progressed to become increasingly efficient in food production and consumption. However, in the process of developing agriculture, man has successfully decreased diversity to shorten and simplify food chains for maximum benefit. Further, in order to make himself the 'terminal' in the food system, he has also tried to eliminate all those organisms which threatened either to compete for his food or go for his life both as predators and parasites. As a result, his enormous numbers today are supported by only a few species of organisms. The relationship between man and such organisms can be regarded as a form of symbiosis emphasizing close mutual interdependence. In developing this relationship, both cultivation and domestication have been highly significant. The evolution has effected cultural modification of man and biological modification of organisms. Consequently such organ-

isms, especially the cultivated plants, have become biologically very different from their wild ancestors of not very long ago. In addition, many of these organisms have, today, become totally dependent on man for their eventual survival and propagation.

Plants in human society

Utilization of plants for more than a mere supply of food began earlier in colder regions. It was first to make fire and then shelter to protect against the cold. Soon plants provided clothing, fats, beverages, boats and other items of our industrial society.

The first development of agriculture may have involved plants with food-storing underground parts. But it is certain that development of cereal did not lag far behind. And once started, it soon spread to various parts of the world. Fig. 81 shows world distribution of three major cereals. Because of many advantages such as high yield per acre, ease of storage, better nutritive value of grain and other multiple uses of the plant, they soon became a necessity in the life of man. All important world civilizations have been known to be associated with one or other of the cereal plants. For example, wheat was the characteristic cereal of most regions of the Old World and maize, likewise, of the New World. In the Asian tropics, however, the cereal that permitted civilization to develop was rice.

There are twenty-four known species of rice; twenty-two being wild and two cultivated. Of the latter, *Oriza glaberrima* is restricted to West Africa and *O. sativa* in Asia, Europe and America. Most probably, the cultivation of rice either originated in India or South-East Asia and then spread to other countries. At present, exchange and introduction of rice varieties of good breed are being conducted on a global scale.

Wheat is another most widely cultivated plant grown in all continents except Antarctica. The grain is said to have been in existence in eastern Iraqi villages some 9,000 years ago. The wheat genus *Triticum* contains a number of both cultivated and wild species. They fall into three groups according to the diploid chromosomes numbers of 14, 28 and 42 present in them. The group with 42 chromosomes comprises the bread wheats (*Triticum aestivum* or *T. sativum* or *T. vulgare*) and is known to have been cultivated in India and central Europe some 4,500 years ago. This group exists only in cultivation and hence might have been the product of unintentional hybridization in some trash heaps of ancient villages.

Similarly, maize is now widely cultivated in all tropical, subtropical and warm temperate regions of the world; it is unknown as a wild plant. This may be attributed to the peculiar structure of its ears with no parallel in the plant kingdom. The grains on the cob are in fact the female flowers with male flowers (tassle) growing on the upper tip of the plant. In its present highly evolved form, maize is an extreme example of an obligate cultigen, being completely helpless without man for propagation. However, the earliest ancestors of modern maize were found in the Tehuacan valley in Mexico with a radiocarbon dating of about 5000 B.C.

In Africa, sorghum and millet seem characteristic of the shifting cultivation, and

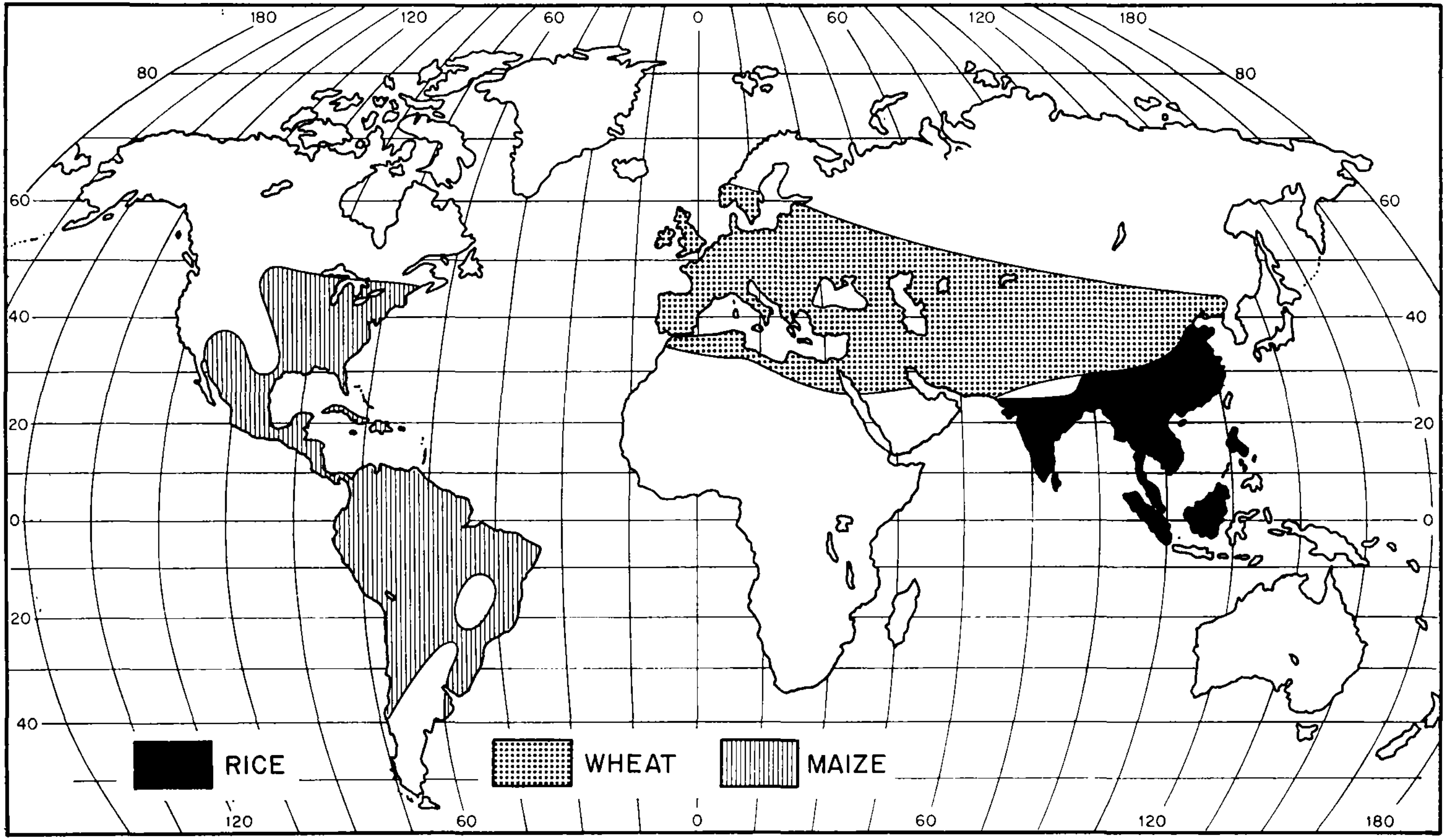


Fig. 81
Three great cereal areas
of the world

Source: J. K. Brierley. *A Natural History of Man*. Rutherford, N.J., Fairleigh Dickinson University Press, 1970, p. 58.

hence they might have been the first to be used by man in or near mountainous regions of the subtropics and tropics.

Other food crops that need mention include cassava, potatoes, squashes, legumes and some vegetables. Of these, legumes, being rich in proteins and fats, are very significant. They are mostly warm season annuals. The characteristic legumes of the New World are several species of *Phaseolus* and *Arachis hypogaea* (peanuts). In general, the Old World has a poorer selection of legumes, the most significant being *Pisum sativum* (peas), *Lens esculenta* (lentils) and *Vicia faba* (broad beans). However, some of the Old World species such as alfalfa have been efficiently developed as forage crops for livestock.

Most oil-producing crops of today are also grown in tropical, subtropical and, to some extent, warm temperate regions of the world. Among these crops, soybeans, sesame, melon, sunflower, cotton-seed, coco-nut and oil palm are most significant. It is expected that such crops may gain increasing importance in the future economy of the newly emergent nations.

The above crops mainly provide the dietary needs of man. Many plants are now cultivated for medicine as well as desirable luxuries. Some of the important drug plants include opium poppy (*Papaver somniferum*) with more than thirty alkaloids, belladonna (*Atropa belladonna*) with atropine as chief alkaloid and Ginseng (*Panax quinquefolium*) with multipurpose use in China. The second category may include sugar, spices and beverages. Sugar-cane, a tall jointed grass, is a native of India, and is now also extensively cultivated in Brazil, and islands of the Caribbean, Hawaii, Java and the Philippines. Beet, developed in Germany at the end of eighteenth century, is an important alternative source of sugar today. The beverages (tea, coffee, cocoa, etc.) that modern man uses are also all products of tropics or near tropics. Major spices are also tropical in origin and their cultivation even today is mainly restricted to such areas. Spices, while adding flavour to food, are also known to enhance the rate of perspiration and hence the cooling of the body.

The last but not least important cultivated plant is the rubber tree (*Hevea brasiliensis*). Despite the fact that it is a native of South America, very little rubber comes from that region. It is now cultivated almost exclusively in South-East Asia, the Congo and western Africa with the major share of the world being produced by Malaysia and Indonesia.

Since most of the cultivated plants are tropical and subtropical in origin, agriculture possibly started very early in South-East Asia with root crops dating between 13000 and 900 B.C., and then spread outward to other regions. From the Tigris and Euphrates rivers, the spread extended with little modification across Europe and to Britain by around 3000 B.C. A similar spread occurred across northern Africa along the Mediterranean.

Table 42. Domesticated mammals of the Old World

Mammal	Wild distribution	Sites and dates of domestication	Geographical expansion	Main uses and aims of selection
Dog	N. temp. Many races (wolf)	Many independent 15000 B.C. (?)	Universal including paleolithic. Australian (dingo)	Hunting Hauling Herding (in domestication of ruminants) Guarding
Reindeer	Subarctic forest and tundra races (E)	N. Eurasia 10000 B.C. (?)	Panarctic	Meat Free ranging, migrating and breeding
Goat (R)	C. Asia and S.E. Europe	Persia, Anatolia 7000 B.C.	Followed both agricultural and pastoral expansions; often displaced by sheep	Meat Milk Skin and hair
Sheep (R)	(as above)	Caspian steppes 6500 B.C.	Main agent of pastoral expansion throughout O.W.	Meat Milk Wool
Cattle (R)	Persia, S. Europe, India (E)	1. Anatolia 6000 B.C. 2. Indus valley 2500 B.C. (independent)	Agricultural and pastoral	Meat (except in India) Milk (except in China) Haulage, everywhere Thrashing corn
Water Buffalo	India (E)	Indus valley 2000 B.C.	To China and Europe with agriculture	(as above)
Yak (R)	Tibet to Mongolia (E)	Nepal c. 1000 B.C.	Confined to arctic climate (over 2,000 m)	Meat, milk, haulage and carriage
Pig	Europe to China (wild boar)	Anatolia 6000 B.C. China? 2000 B.C.	Originally woodland; later near universal with agriculture	Meat and skin (and collecting truffles)
Ass	N. African steppes	Egypt 4000 B.C.	With pastoralism (Babylonia 2000 B.C.)	Carriage Riding Milk
Onager	S.W. African steppes	Sumeria 3000 B.C.	(E) displaced by horse 1500 B.C.	Haulage
Horse	Eurasian steppes tundra and forest	Caspian 2250 B.C.	With pastoralism, to Egypt 1600 B.C., to China 1500 B.C.	Meat Haulage Riding especially in war Milk
Camel (R)				
a. Dromedary (one-hump)	Arabia (E)	1200 B.C.	N.W. Africa to India (southern)	Carriage Riding Milk
b. Bactrian (two-hump)	C. Asia	500 B.C.	Anatolia to Mongolia (northern)	(as above)
Elephant				
a. Indian	Syria, India, China	Indus 2500 B.C.	S. Asia	Jungle haulage Battle
b. N. African	N. Africa (E)	Egypt 280 B.C.	Italy (Hannibal)	Battle only (E)

R: Ruminant
E: Extinct as wild or as domesticated animals
O.W.: Old World
N.W.: New World

Source: J. K. Brierley. *A Natural History of Man*. Rutherford, N.J., Fairleigh Dickinson University Press, 1970, p. 60–1.

Animals in human society

In contrast to large-scale cultivation of plants, only a few species of animals were domesticated. This is evident from the Mayan and Incan civilizations of America. But as in the case of plants, early usage of animals was mostly for food (meat or milk). Soon their bones were used as tools and ornaments and their skins as clothes. Some animals such as cattle, donkeys, horses, camels and elephants were also used for transporting man and his goods. Slowly man's relationships with other animals multiplied. As he constructed shelters, a wide variety of animals moved in to share his house and habitat alike. This is true even today; many lizards, mice, scorpions, cockroaches and a host of others are found closely associated with our dwellings.

Despite plant cultivation, domestication of animals proceeded rather slowly at first, but by about 2000 B.C. animals were known to have been domesticated. Numerous mammals of Europe, Asia and Africa are listed in Table 42 along with information concerning their origin, domestication, expansion and uses. Among the few animals domesticated early in tropical America are the alpaca, llama and guinea pig.

The dog is regarded as one of the earliest animal to be domesticated, perhaps as far back as 15000 B.C. But how, and where this came about and what particular species was involved are all matters of controversy. Some believe it started as a pet while others assert that it adopted man and followed him in his hunting expeditions, living as it were on 'left-over crumbs' from his meal.

Gradually, with the development of agriculture, domesticated animals were used increasingly for performing agricultural tasks. While this greatly increased food-producing capacity, it also required diversion of substantial proportions of food for maintenance and further domestication of such animals for additional purposes. Table 43 summarizes current objectives of domestication.

Table 43. Modern purposes and examples of domestication

Meat	Milk	Eggs	Hunting	Labour	Skins, fur, etc.	Pets
Ox, sheep, pig, goat, rabbit, chicken, guinea, swan, goose, turkey pheasant, duck, pigeon, peacock.	Cattle goat, ass, camel.	Chicken, goose, duck, guinea, turkey.	Dog, horse elephant, cheetah, hunting leopard, hunting hawk, ferret.	Horse, ass, ox, mule, water buffalo, elephant, camel, reindeer, dog, yak.	Horse, ox, sheep pig, goat, dog rabbit, sheep, vicuna, alpaca, fox, skunk cattle, mink.	Dog, cat, horse, rabbit, monkey, guinea pig, hamster, canary, peacock, pheasant, swan, parrot, and others.

Impact of domestication and cultivation

In developing agriculture, to suit his needs man has gradually succeeded in replacing more stable biological communities with less stable man-made ones. For example, a wheat field can be reaped with ease but it also provides a most suitable habitat for wheat-loving insects and fungi which, if not checked, may prove catastrophic to the entire crop. Some of the introduced crops of tropical Africa have insect pests that lower the yield considerably. They are both introduced and local, the latter simply taking advantage of new opportunities.

Similarly, the creation of national parks in eastern and central Africa, primarily for conservation of native ungulates and tourist attraction, has created special problems. Through over-grazing, elephants have caused destruction and retreat of the forest in some areas. In many places, man has destroyed many forests to claim timber and land for housing and agriculture. Hence, there is a great need on our part to balance efficiency against safety and stability.

Pest control

Man's demands for additional food, fibre and timber, etc. have resulted in greater use of land and water with important alterations in the habitats of wild animals. As a result, many species have suffered large population losses even to the extent of their extinction, and yet others, finding these changes desirable, have multiplied so much that problems are created which adversely affect man's health and general welfare, his crop production and his forests. Development of agriculture has greatly aided natural selection for such organisms to develop. Furthermore, in moving materials around the world, man has helped many organisms to find suitable habitats devoid of natural parasites and predators. Thus, in the absence of natural checks, they multiply to the limit of their food supply to become major pests, interfering and competing with man's efforts at food production. Such organisms are not wanted and man must wage constant 'warfare' against them to protect what he wants and needs for his own use. Weeds may be pulled out by hand or removed by cultivation. Animal pests are much harder to control. Of the various kinds of nuisance, destructive, disease-carrying, or predatory pests, insects and rodents are the main targets of control. Birds present an equally important problem in this respect. Of lesser significance are rabbits and ungulates (deer, goat and antelope), which cause local plant damage and compete with livestock for forage. Thus pest-control measures have been developed. Increasingly, emphasis has been placed on chemical poisons (pesticides) as weapons in this struggle. The subject has been treated previously in this work. Here it may be mentioned that the use of pesticides is not without dangers to man himself. Controversy has especially centred around DDT and other related substances. They are toxic not only to insects but to men and a variety of other organisms. Nevertheless, it is sometimes highly desirable in a situation where pests pose a greater threat to human health and crops.

In nature, however, biological control seems more successful than a chemical one. This control involves the use of natural predators, parasites or pathogens of pests to keep their populations in check. Lady beetles, for example, are effective as predators upon numerous agricultural pests, and calcid wasps are parasitic upon various pests. This method is especially desirable because it is safe, permanent and economical. But the use of such parasites and predators requires a thorough understanding of the relationships among the organisms and their environment if the effort is to be effective and new imbalances are to be avoided. If such relationships, holding pest populations in a delicate and dynamic balance in the ecosystem, are inadequately understood, an organism introduced as a biological control may itself become a pest as it feeds on an unanticipated food supply. For instance, the mongoose, which was introduced into some areas to control populations of snakes, has itself become a pest by destroying chickens. In fact, unilateral use of any control, including even biological control, can have deleterious and unintended side effects on an agricultural system. A new variety of plant may prove more susceptible to fungus disease occurring in the area. Similarly the reduction of a pest through a biological agent may disturb the balance and relationship of other organisms that feed on it or may otherwise result in modification of the environment. In Australia *Aphidius* was introduced to control the carrot aphid, *Cavariella aegopodii*. This resulted in the virtual extermination of carrot motley dwarf virus. Such examples are numerous in the literature.

Consequently, integrated control involving pest population management is now highly recommended. It aims at utilization of all suitable techniques and procedures (including use of insect-resistant varieties of plants) either to reduce pest populations and maintain them at levels below those causing economic harm or to so manipulate the populations so that they are prevented from causing even greater harm. The underlying theme is harmony rather than disruption, and the strategy of development of such a control must be so patterned as to aim for profitable production and minimum interference with nature's relationships.

Climax and biomes

In the absence of any disruption or interference, a natural biotic community, through various successions, eventually reaches a stage that is dynamically more harmonious and stable. It becomes self-perpetuating and may persist for centuries so long as environmental factors remain essentially constant. At this stage it is called the 'climax community'.

There are no distinct boundaries between biotic communities. Like the environmental conditions (micro-climate, wind, temperature and physiognomic features) they change continuously both in space and time. Abrupt changes are rare. However, cumulative effects could be large enough to separate different communities and the climax, therefore, can be meaningfully considered only in relation to a local or geographical set of environmental conditions. Thus, a number of major climax formations or 'biomes' as shown in Fig. 63 can be recognized. Some biologists prefer to use

terms such as major ecosystems, provinces, biochores and regions to describe them.

The distribution of biomes, especially in the Old World, corresponds with the latitudinal lines. Understanding of the pattern, however, seems to be complicated by interactions between climate, soil and vegetation, each influencing others and being influenced by them. As such, each biome is quite variable and includes many different local communities. Of the world's major biomes, only tundra and taiga (northern conifer forest) form circumglobal belts. The biomes south of taiga experience more variation in the amount of rainfall. As a result these biomes are isolated in different biogeographical regions, having ecologically equivalent but often taxonomically different species.

The tundra

The tundra lies north of latitude 57° N, forming a circumpolar band in North America, Europe and Asia with a narrow interruption by the North Atlantic and the Bering Sea. It is characterized by the arctic climate with dark winter and daylight summer. The subsoil is permanently frozen (permafrost). The permafrost line is the ultimate limit of plant root growth.

The vegetation consists of a relatively few species of mosses, lichens, grasses, sedges and dwarf woody plants, the tall trees being absent. Of these mosses especially sphagnum and lichens (*Cladonia*, reindeer 'moss') cover larger areas while the low flowering herbs characterize the marshes and the poorly drained areas. Tundra plants most of the time remain in a state of suspended animation, coming to life only in the short growing seasons of summer thaws.

The animal life includes mainly birds and mammals which seem well protected by their feathers and furs respectively. Of the birds, many, especially the shore birds (sandpipers, plovers) and waterfowls (ducks, geese), are only summer visitors. Among the mammals, musk oxen and caribou (wild reindeer) are large herbivores living on moss and lichens while arctic wolves and foxes are carnivores preying on arctic hares and lemmings. Polar bears are also common, frequenting coasts and ice floes. Among the invertebrates, insects, particularly flies and mosquitoes, are most numerous. Their eggs and larvae are resistant to cold and adults appear in summer to the extent of becoming major pests in the area.

Taiga (coniferous forest)

Taiga occurs between 45–57° N latitudes as a broad belt of evergreen forest south of tundra in both North America and Eurasia. Extensions of the zone occur even in tropical mountains. Like tundra, this biome is not represented in the southern hemisphere and is characterized by many bogs, ponds, etc. Again, like tundra it has severe winter temperatures, but summers on the other hand are warmer, longer and characterized by greater precipitation, encouraging growth of heavy and hardy trees. The soils are deficient in minerals such as calcium, nitrogen and potassium because of their leaching through large-scale water movement in the region.

The total number of species in taiga is larger than in tundra, but smaller than those lying farther south. The characteristic vegetation is of coniferous trees such as spruce, fir and tamarack. Among these spruce is predominant. Alder, birch and juniper are also common. In dry weather some vegetation often gets burned. In such areas aspens and birches appear as pioneers to be succeeded by conifers later on.

Of vertebrate fauna, the moose (called 'elk' in Eurasia) is the most conspicuous. Other common mammals of the zone include black bear, wolves, lynx, wolverines, marten, squirrels. Birds are common but they are mostly summer breeders and migrate southward in the fall. Coniferous seeds provide food for many small animals such as siskins, squirrels and cross-bills.

Seasonal periodicity is conspicuous, the snowshoe hare-lynx cycles being the classic examples. Among invertebrates, the insects are most numerous and lie dormant during the severe winter, but on the return of the breeding season there are outbreaks, especially of bark beetles and defoliating insects, e.g. saw flies and budworms.

Temperate deciduous forests

These forests occur in most of eastern United States. Originally, however, they covered the British Isles, the whole of central Europe, parts of China, Japan and Australia, and the southern tip of South America.

The climate is characterized by distinct cold winters and warm summers, and with moderate and evenly distributed rainfall (75–150 cm). The winter half is the period of nearly suspended animation with trees and shrubs bare. The summer half, on the other hand, is the period of activity with plants going through their complete life cycle.

The deciduous forests include many more species than do the first two biomes and these are arranged in subdivisions, each including two or more common species characterizing the group as in the beach-maple, oak-hickory and elm-ash-maple climax forest types. In all these, herbs and shrubs form distinct storeys. Food in the form of pulpy fruits and nuts is plentiful.

Common herbivores include deer (north America and Eurasia), wild pigs (Eurasia). The main predators are the entire cat family. *Felix concolor* (puma or mountain lion), which occurred both in North and South America, is now extinct in the eastern forest. Squirrels are plentiful in trees and so are the tree-nesting birds, especially woodpeckers and ovenbirds. The soil biota is also well developed with fungi and invertebrates swarming in the leaf- and mould-covered forest floor.

Evergreen subtropical forest

Such forests occur in most of eastern United States, most of central Europe and a part of eastern Asia, where the moisture content is high but temperature differences between summer and winter are not pronounced.

The climax communities are predominantly composed of broad-leaved evergreen

trees ranging from oaks, magnolias and hollies in the more temperate regions, and figs and wild tamarinds in the more tropical regions. Coniferous vegetation (white pine, hemlock, cedar) may be predominant in some locations while palms (*Sabal palmetotto*) may be prominent in others. There is a diversified understory of shrubs and herbs.

Tropical rain forest

This biome is represented in three main zones: South America (the Amazon and Orinoco basins) and Central American Isthmus; Africa (the Congo, Niger and Zambesi basins) and Madagascar; Indo-Malay (Malaysia, Indonesia, etc.) and New Guinea regions. Life in this biome is most exuberant and diversified. In contrast to a temperate forest with two or three dominant species (ten at best), a tropical rain forest may have 100 or more species (500 have been counted in one such forest).

The dominant trees are very tall, usually 20–40 metres in different situations, and their tops form dense canopies which intercept both light and rain. They also shield the forest floor from wind and hence reduce the rate of evaporation underneath. Thus conditions of life and micro-environments are different at different elevations resulting in a distinct vertical stratification of fauna and flora.

The floor is almost constantly dark, humid and with little change in temperature (around 25° C). Among green plants only a few with meagre photosynthetic requirements may grow there. In addition to tall trees most vegetation include epiphytes and liannas.

Most animals are nocturnal and those active during the day live in the upper strata of vegetation in contrast to temperate forests where most life is near the ground level. In addition to arboreal mammals (monkey, lemurs, flying foxes), there is an abundance of rodents, squirrels, iguanas, geckos, snakes, birds and frogs. Ants, grasshoppers, termites, flies, butterflies, beetles and other insects are equally abundant and many form symbiotic associations with epiphytes. Ground-dwelling herbivores include musk deer, wild pigs, etc.

Fruits and termites form staple food for many animals.

Grasslands

They are variously called prairies, steppes, savannahs or pampas in different parts of the world but because of ecological similarities between them they may be grouped as grasslands. They occur in the drier interiors of continents in both tropical and temperate zones where rainfall is about 25–75 cm per year; not enough to support a forest but more than sufficient for the growth of grasses and small herbs.

The tropical grasslands (often called savannahs) are found in Africa, South America and Australia, while their temperate counterparts are mainly located in western United States, Argentina, Australia, southern Russia and Siberia. Temperate and tropical grasslands are remarkably similar in appearance, in spite of the difference in species they support. Trees and shrubs may be found in grasslands either scattered

individually or in narrow belts along rivers and streams. The significance of grasslands lies in the fact that our agricultural crops of corn and wheat have been developed from grasses by artificial selection.

The grasslands provide natural pastures for cattle, sheep and goat and their rich soil is well adapted for the cultivation of wheat and corn, but under consistent overgrazing and overplowing they can be turned easily into man-made deserts.

Primary consumers are very conspicuous and numerous in grasslands and include bisons and pronghorns in the temperate regions, and hippopotamuses, zebras and several species of antelopes, etc. in the tropical regions. In addition, hares and rodents are also common. Australian counterparts of these animals are kangaroos and small, pouched 'mice'. Predators are equally significant and include wild dogs, lions, etc. preying on the ungulates; and snakes, weasels, etc. on the smaller herbivores.

Voracious insects such as grasshoppers and locusts feeding on vegetation are also very numerous.

African savannah is now posing a problem in land-use for raising the quality and quantity of food for Africa's rising population. In view of the discussion earlier in the chapter it may be sometimes better to develop and harvest the native herbivores (antelopes and wild beasts) on a sustained-yield basis rather than try to rear cattle. The introduced species, besides being more susceptible to tropical parasites and diseases, may also disturb ecological balance which may be difficult to restore.

Deserts

The Sahara in Africa is the largest desert biome in the world. Other major deserts include the Australian Desert, the Turkestan Desert (U.S.S.R.), the Atacama-Peruvian Desert (Chile and Peru), the Patagonian Desert (Argentina), the Thar Desert (India and Pakistan), the Gobi Desert (Mongolia), the Arabian Desert, etc.

Deserts have very low rainfall (from 50 to 100 mm per year) which often comes at erratic intervals. They are further characterized by extreme daily temperature fluctuations.

Only some deserts such as parts of the Sahara are almost barren, but more commonly the vegetation consists of scattered and drought-resistant shrubs (e.g. sage-brush and greasewood) and succulent plants capable of storing much water in their tissue (e.g. cacti and euphorbias). Most other plants are annual herbs capable of blooming and reproducing seeds all within a few days. In brief, plants are so modified as to conserve water supply.

Most animals are primarily nocturnal so as to avoid intense heat during the day. They are also adapted to the scarcity of water. Except for the camel large mammals are rare but small rodents (e.g. kangaroo rats) are numerous. Other common animals are snakes, lizards, a few birds, arachnids and insects.

Biogeographical considerations

Introduction

All organisms are not found uniformly distributed in all parts of the world. They are not even found everywhere that they could survive. Cacti, for example, are found in American deserts and not in African deserts. Likewise, central Africa has elephants, lions and gorillas while Brazil, with similar climate and environmental conditions, has prehensile-tailed monkeys, sloths and tapirs.

There are two lines of explanation for such geographical distributions. According to one it is limited by suitable environment and climate. The presence of cacti in deserts and its absence from wet forests in America can be explained on this principle. As such, distribution of biomes both along latitudinal and altitudinal gradients can be illustrated (Fig. 82). Temperature and solar radiation have a distinct north-south gradient which, along with precipitation, affect vegetation distribution. Tropical lowlands are without frost but with intense sunshine and little variation in seasonal temperature. As one proceeds northward from the equator, seasonal variations of temperature become more pronounced, frost common through longer and longer parts of the year and summer duration reduced. Corresponding altitudinal situations are encountered if one climbs a very high mountain. Consequently, conditions on a mountain top are similar to those at sea level farther north. Similarly, parallel zoning of biotas is encountered both horizontally and vertically, as shown in Fig. 82.

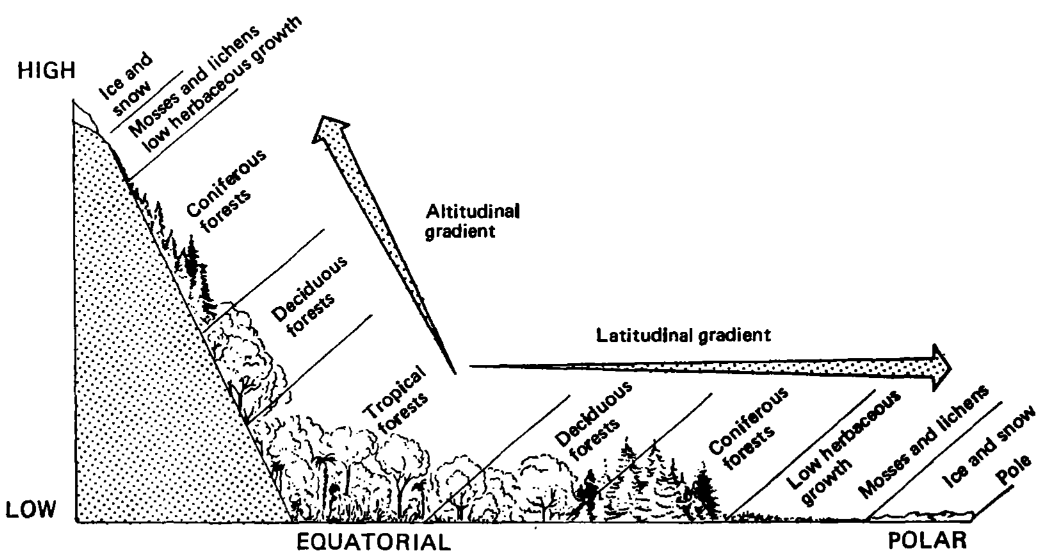


Fig. 82
The parallel between the horizontal and vertical distributions of life zones

Source: G. G. Simpson; W. S. Beck. *Life: an Introduction to Biology*. New York, Harcourt, Brace & World Inc., 1965, p. 711.

The second view maintains that the above explanation is only partially true. The fact that deserts occur in Africa, but cacti do not, needs further understanding in terms of the evolutionary history of each species. Organisms arise in a particular region and reproduce to multiply. But since the reproductive capacity always exceeds the carrying capacity of the environment, there is always pressure within a population to expand its niche or to spread into new territories.

Dispersal and distribution of species

There are at least three major conditions which must be met before a species can successfully spread and establish into a new area.

Firstly the species must possess the physiological potential to survive and reproduce in the new area. It simply means that the organism must be able to adjust to the 'new home' and its sources of food. Once colonized, new evolutionary changes may follow under selection pressure. For example, the camel family originated in North America. Later, some camels spread to South America where they evolved into llamas, guanacos, alpacas, etc. while others reached the Old World (via the Alaskan-Siberian land bridge) and evolved into Bactrian camels and dromedaries. In both cases they still survive even though they have become extinct in their original homeland of North America. This shows that migrating camels had evolved some characteristics in their original habitat which served as premium for further evolution and successful colonization in new homes.

Secondly, the colonizing species must not, at the very outset, encounter severe competition in the new habitat. It must, in other words, find some suitable niche that it can fill there. Briefly, the species must have the ecological opportunity to establish itself.

Thirdly, the dispersal is often conditioned by some continuity of suitable habitat. Despite the first two conditions an organism cannot colonize new areas if it has no physical access to them. The spread of cacti from America to Africa, for example, seems to have been blocked by the barrier of the Atlantic Ocean.

All organisms have some means of dispersal. In most animals, active locomotion (flying, walking, swimming) is most obvious. Even many sedentary marine animals have a free swimming larval stage. Thus, all these organisms constantly change their geographic positions though the shift from the point of origin may not be large in one's own lifetime. But over many generations the cumulative effect may be considerable. However, it must be distinguished from migrations which involve periodic movements of whole populations among regions previously known.

For small organisms (seeds, fruits, insects, spiders) passive transport by wind, water and other animals is the chief mechanism of dispersal. Man has intentionally or unintentionally spread many plants and animals from one place to another. The recent history of Krakatoa, a small volcanic island in the East Indies, illustrates beautifully a natural experiment in dispersal and colonization of a new habitat by both animals and plants. Since its destruction and annihilation of all life from it in 1883, the island is now again alive with organisms from the next closest island about twenty-five miles

away, these organisms having dispersed there by various agencies mentioned above.

Dispersal does not depend on means alone. There must also be some route along which it may take place. Fig. 83 shows world dispersal routes with barriers, filters and corridors.

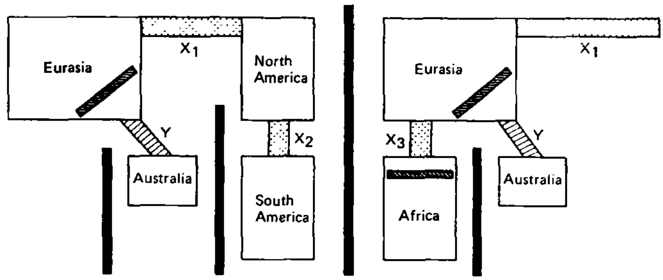


Fig. 83
World dispersal routes: barriers, filters and corridors
Source: G. G. Simpson; W. S. Beck. *Life: an Introduction to Biology*. New York, Harcourt, Brace & World Inc., 1965, p. 737.

X_1, X_2, X_3, variable filter bridges or corridors, Y, variable sweepstakes route ||||||||, constant barrier during the Age of Mammals; /////////, somewhat variable land barrier.

In oceans the main routes of dispersal are the currents. Planktonic distribution is strongly influenced by them. The nektonic animals, however, are affected by ecological barriers (temperature, salinity, etc.) in their dispersal.

On land, the zones of prevailing westerly winds are the major routes from west to east for plants and animals with air-borne dispersal phase. In view of greater similarities between natural communities of China and Europe, many plants and animals must have traversed the great dispersal route between western Europe and northern China. However, for every organism each route offers a certain probability of dispersal. For example, an ocean almost completely prohibits movement of horses or elephants but water-resistant coco-nut palm seeds may cross it in large numbers. Similarly, some forest animals may have difficulty in crossing grasslands while others may do so freely. In brief, that which is a barrier to dispersal for one species may be an easily negotiable path for another, with gradations in between. As such a route is a 'corridor' if it allows most species of a biota to disperse (e.g. the Eurasian route mentioned above). It is called a 'filter' when it passes parts of biotas and holds back the rest (connexion between North and South America). If on the other hand the chances of crossing from one region to another are almost negligible, the intervening zone is called a 'barrier'. An ecological barrier may be as potent as the physical one. Plant seeds may drift away for hundreds of kilometres or may even cross an ocean but if they land on an unsuitable soil in an unsuitable climate, they will not grow. Chances of development are equally slim even if they find suitable abiotic environment amidst a well-established, hostile and competing biotic community. Faunal regions are often delimited by major barriers. For example, the Himalayas and other mountains are the barriers between the Palaearctic and the Oriental regions in Asia, and the Sahara separates the Palaearctic and the Ethiopian regions in Africa.

However, there are sometimes exceptional cases when populations cannot spread across barriers by the usual process of expansion or migration but they do so as if in

one jump. This is called 'sweepstakes dispersal' as in the colonization of Krakatoa island. The possibility of sweepstakes dispersal is much higher for plants, especially those with wind-borne spores or seeds. It is amply supported by data on the Pacific islands and elsewhere.

Biogeographical realms and associated fauna

Despite all the present knowledge, we still cannot fully explain the distribution of all animals and plants in the world.

As indicated earlier we need to consider the historical element as well. The earth itself and the organisms on it are constantly changing and present distributions are in very large part the result of past conditions. Hence to understand the present geography of life we must combine present situations with the evidence from fossil records, past geological formations and connexions, and the climate of the earth's land masses and

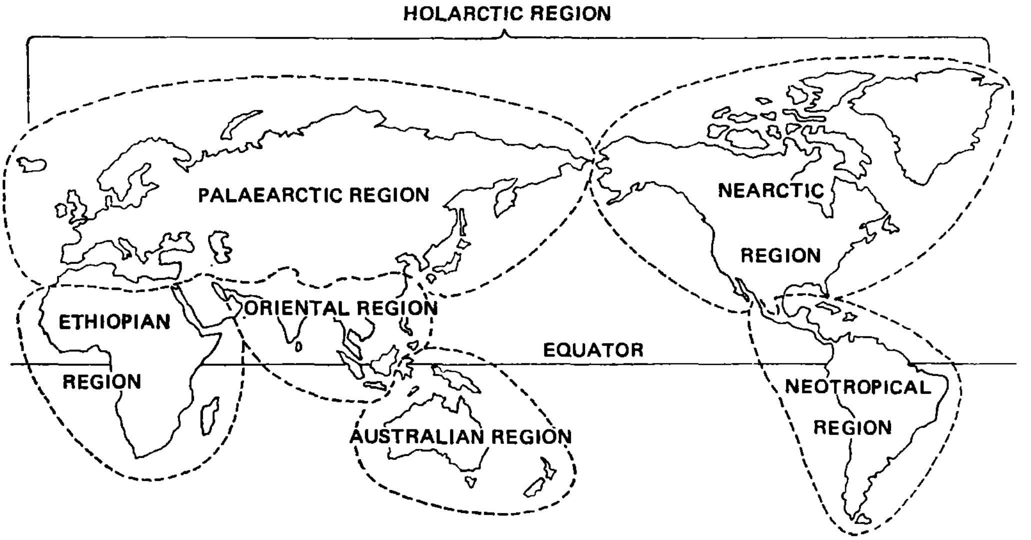


Fig. 84
Biogeographical realms (based mainly on the distribution of mammals)
Source: G. G. Simpson; W. S. Beck. *Life: an Introduction to Biology*. New York, Harcourt, Brace & World Inc., 1965, p. 728.

oceans. Thus, careful study of distribution of plants and animals indicate the existence of six major biogeographical regions or realms as described below and shown in Fig. 84. Each is characterized by certain unique organisms. Although the divisions were originally based on distribution of land mammals, they have since been found valid for many other groups of organisms.

The Australian realm. This includes Australia, New Zealand, New Guinea and other islands of the East Indies. Wallace's line separates it from the Oriental realm. The line also passes between the islands of Bali and Lombok. These islands, though separated by a narrow channel about 32 km wide, show fauna and flora which are much different.

The biota of Australia is unique. Monotremes, the egg-laying mammals (duck-billed platypus) and many marsupials are found only here and no where else. Other remarkable animals of Australia include running birds (emu, cassowary), the lyre-bird and cockatoo and Dipnoan fish *Ceratodus*. Equally interesting is the fact that many species of placental mammals, widespread in the rest of the world, were conspicuously absent from the continent until at the time of European discovery. Even in the case of the dingo (native Australian dog) and wild pig, the human influence cannot be excluded.

Evidence both biological and geological indicates that Australia is the only southern continent which had no direct link with the northern hemisphere land mass for a very long time. However, this isolation, as shown by the fossils range of *Ceratodus* and *Eucalyptus*, did not always prevail. Moreover, most of the western islands of the East Indies were possibly interconnected as an extension of the Asian land mass several times in the past. As such, ancestors of most if not all organisms now living in the Australian realm might have migrated there by island-hopping over millions of years in the remote past.

Among the earliest arrivals in Australia were apparently the marsupials. Finding no competition from placental mammals there, they spread over the whole continent filling niches similar to those occupied by placentals in the rest of the world.

Later, rats (placental rodents) also reached Australia from the world continent. They are now numerous there.

The Neotropical realm. This consists of South America, Central America, southern Mexico and the islands of the West Indies. The characteristic fauna includes alpacas, llamas, prehensile-tailed monkeys, blood-sucking bats, sloths tapirs, ant-eaters and many birds species (toucans, puff birds, etc.).

Like Australia, South America has been an island continent unconnected to other major land masses through the greater part of its history. However, as indicated by more varied stock of land mammals including both marsupials and placentals, it must have had a connexion with North America early in the age of mammals. Later, the land bridge with North America was broken and the mammals evolved in isolation in South America. Even during the period of isolation lasting about 60 million years, a few additional placental mammals, especially the ancestors of modern New World monkeys and rodents, are said to have crossed from North America at points where water barriers were not insurmountable. On the re-establishment of the bridge, there were more migrations under which South American species such as the opossum, armadillo and porcupine crossed over to North America. It is said that the bridge only served as a filter route rather than a corridor for general dispersal.

The Palaearctic realm. This covers Europe, Africa (north of the Sahara Desert), Asia (north of the Himalayas), Nan-Ling mountain, Japan, Iceland and the Azores islands. Important native animals are moles, deer, oxen, sheep, goats, robins and magpies.

The Nearctic realm. It comprises only Greenland and North America (up to Mexico in the south). Many of the Palaearctic species are found here. In addition, prairie dogs, opossums, skunks, racoons and turkey buzzards, etc. are indiginous to the region.

Because of the many similarities in fauna, the Palaearctic and the Nearctic are grouped together in one realm called the Holarctic region.

The Oriental realm. This includes India, Sri Lanka, Indo-China, southern China and Malaysia. It is known for its tigers, elephants, gibbons, orangutans, black panthers and tarsiers.

The Ethiopian realm. This covers a part of Africa (south of the Sahara Desert), and Madagascar island. The gorilla, chimpanzee, zebra, rhinoceros, hippopotamus, giraffe and many birds and reptiles are native of the region.

The four biogeographical regions (Palaearctic, Nearctic, Oriental and Ethiopian) covering Europe, Asia, Africa and North America, formed a relatively continuous land mass, often called the 'world continent' through most of the geological time. The obvious barrier between the Palaearctic and the Ethiopian realms is not the Mediterranean but the Sahara Desert in Africa. The Himalayas similarly separate the Oriental region (tropical Asia) from the Palaearctic northern Asia.

There are problems of resemblances and differences between these regions. For example, North America is connected to South America but its fauna is more like that of northern Asia. Similarly, northern Africa, though not connected directly to Europe, has fauna with a predominantly European element. Again, fauna of central and southern Africa resembles more that of distant southern Asia rather than that of northern Africa. Many other problems of biogeography arise from disjunctive distributions in which two closely related groups of organisms occur in widely separated regions with no closely related forms between, as in the distribution of tapirs. One conclusion, however, is certain: the two disjunctive species must have had a common ancestor. But how and through what route they dispersed are difficult to ascertain.

9 The oceans

At sea ecosystems have hardly any boundaries at all. In fact, the entire Indian ocean can be classified as an ecosystem. Yet, it is in communication with both the Atlantic and Pacific oceans which illustrates the immensity of the marine environment as an ecosystem. In the sea there is life throughout the entire water column, despite the fact that light, as seen by human eyes, only penetrates the water down to a depth of 579 m. However, this does not mean that photosynthesis occurs at such a depth. Conservative estimates place the lower photosynthetic level at about 200 m.

But the products from life organisms of the oceans' upper levels are so rich that they, through marine food webs and nutrients washed out from land, sustain all oceanic life from the surface down through the pelagic region to the benthic region. Many organisms rise cyclically to the surface and upper water levels to feed actively on the food available there; others are nourished by decomposition products which, like continuous rain, drop down from the upper sea levels. Moreover, ocean currents sweep at all levels, often teeming with life organisms. In the oceans there is throughout a trophic structure in contrast to terrestrial habitats within the atmosphere which are poor in organic life a few metres below the surface of the soil, or above about 3,000 m of this surface, where the wind carries relatively few living things. This means that marine ecosystems are larger and more complex than terrestrial ones. They are also characterized by a high reproductive rate. The phytoplankton is as a primary producer the basis of marine food chains and outnumbers considerably rooted plants of the sea, but in turn is itself outweighed by zooplankton. This fact emphasizes another oceanic feature of great ecological importance: the mobility of marine organisms. Not only is the plankton mobile by its vertical movements night and day, but it can also be transported horizontally by currents to wide areas in the ocean, where they nourish secondary consumers from fish to whales. Another characteristic feature is that in the oceans the pyramid of life is not broadest at its base as on land.

When discussing the oceans one has to bear in mind that they cover 71 per cent of the earth's surface. If they were evenly distributed, they would cover the entire earth about two miles deep.

Physical and chemical conditions

Ocean waters are oxygenated by waves and in coastal areas by rivers but the role of green plants, for instance algae, is perhaps even more important, in this respect. The oxygen content of the sea-water is by no means uniformly distributed. It varies between 0 and 8 ml/l, more often within the range 1–6. Since the gas is more soluble in cold water than in warm, it is generally greater at high latitudes than nearer the equator. An oxygen-minimum layer usually occurs at a depth of 400 to 1,000 metres, being most evident in low latitudes where the water at 400–500 m is often found almost completely lacking in oxygen.

The salinity of the ocean surface is variable depending on a number of factors: climatic and geographic differences, influx of freshwater from rivers and melting ice; ocean currents, differential evaporation at various latitudes, and so on. Tropical seas have higher salinity than temperate ones, owing to greater evaporation. Below a depth of about 300 m the ocean's salinity is almost constantly about 3.5 per cent. The Mediterranean and Red seas are subjected to special conditions, because they are connected with the adjacent oceans by more or less shallow straits; they have a higher salt content than the Atlantic or Indian oceans due to more rapid evaporation and smaller influx of freshwater. The salinity in the Mediterranean rises to 4 per cent and in the Red Sea to 4.7 per cent. The tropical coastal waters including those of the straits of Malacca and Singapore have lower salinity varying around 3.1 per cent.

The salts consist of chlorides of sodium, magnesium, calcium and potassium of which common salt ($NaCl$) is the most common. Sea-water also contains carbonates. Nitrates and phosphates exist in small quantities but are of great importance in the chemical ecology of the ocean.

Physically the currents play a considerable ecological role in the ocean. Due to pronounced differences in water temperature in tropical waters and arctic waters, there are also great differences in density. This phenomenon (and also the rotation of the earth) creates currents which move in vertical-horizontal circulation between the tropics and the polar regions. The warm tropical water moves on the surface towards the Arctic and Antarctic regions, where it gradually cools, sinks and starts moving slowly as an undercurrent back to tropical sea areas, where it is successively heated and replaces the warm surface water which is again flowing to colder areas.

This oceanic circulation points out the global dimension level of the marine ecosystem, but of course continental land masses interfere in many ways in the direction of the oceanic currents, making them very complex.

However, as a general pattern the surface water forming the great North and South Equatorial currents flow from east to west in the Atlantic, Indian and Pacific oceans. In the Atlantic, warm waters from the tropics off Africa move first in north-western and western directions (deflected by land masses and Coriolis force) towards South America and the Gulf of Mexico, from where they proceed under the name of the Gulf Stream towards the Arctic areas of the northern Atlantic. The surface movements of the Pacific Ocean show a broadly similar pattern to those of the Atlantic. The Kuro Siwo current, flowing in a north-easterly direction past the south

island of Japan, is the counterpart of the Gulf Stream in the Atlantic. A cold current, the Oyo Siwo, flows down the western side of the Pacific towards the north island of Japan. In the northern part of the Indian Ocean, the surface circulation is complicated by seasonal changes in the direction of the monsoons. In the South Indian and South Pacific oceans there is a counter-clockwise surface gyre and a deflection of water into the Southern ocean such as occurs in the South Atlantic. Upwellings of deep waters (100–200 m) supply nutrients to the Canaries current, Benguela current, Peru current, California current and West Australian currents and hence these are all areas of high fertility. The phenomenon also occurs in areas such as the Equator where currents

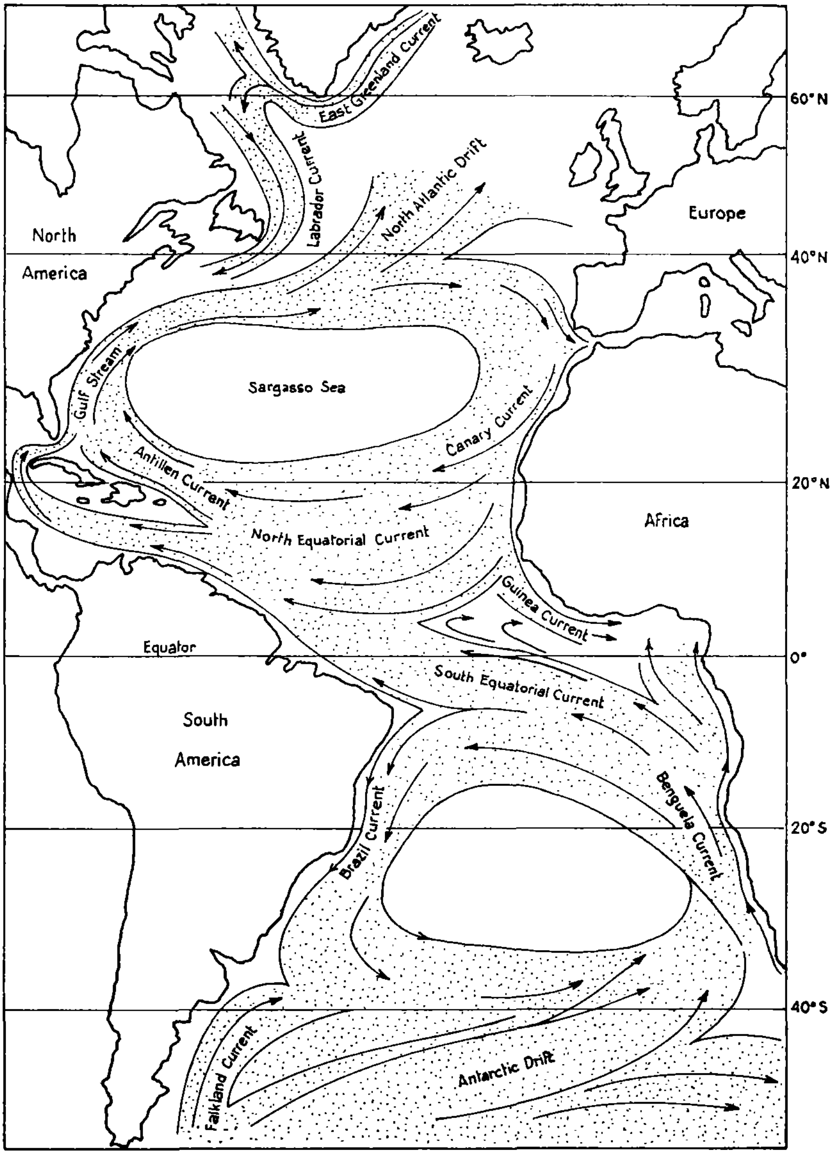


Fig. 85
Major currents of the Atlantic Ocean

and counter-currents diverge and is thus significant in maintaining the productivity of tropical waters. (See Fig. 85 for major currents of the Atlantic Ocean.)

Temperature is another physical factor of great ecological importance in the ocean as well as on land. Most strong currents in tropical and temperate seas move in the upper layers from the surface down to about 100 m. This vertical zone is also influenced by water movements stirred by winds. Beneath this surface zone is a transition layer, a so-called thermocline, in which temperature decreases rapidly with depth. This transitional thermocline separates the warm waters of the surface layers from the cold waters of the ocean depths, where the water temperature is almost constantly cold, reaching about 5–2° C at the bottom. This immense horizontal as well as vertical cold-water zone is connected with the cold-water masses of the polar areas. Hence in the tropics there are in the deeper oceanic waters temperature conditions similar to those in the Arctic and the Antarctic, but nevertheless life in these deep-water areas differs considerably due to the various organic matter dropping down from the upper layers which harbour totally different organisms in tropical and polar waters.

The continual circulation of the oceans and their enormous heat capacity help to maintain a general low temperature variation in the sea despite geographical and seasonal differences in absorption and radiation of heat.

The greatest seasonal variation in temperature (8° to 20° C) have been recorded in the China Sea and the Black Sea. The highest temperatures, however, are found at the surface in low latitudes where much of the oceanic surface water is between 26° and 30° C. In the Persian Gulf, which is shallow and partially enclosed, the surface temperature as high as 35° C has been recorded.

The Red Sea as far as temperature is concerned is quite peculiar. In the middle part of the Red Sea the mean surface temperature is 35.6° C. Below the thermocline, between a depth of 700 m and 2,190 m (greatest depth of the Red Sea) the temperature is as exceptionally high as 21.5° C, whereas at the same depth in the Indian Ocean, the average temperature is only 2.11° C. Temperature influences rate of photosynthesis directly and hence the marine productivity. The production is also affected indirectly through its effects on movement and mixing of water and the subsequent supply of nutrients to the euphotic zones.

Another physical factor of importance to oceanic life is pressure. Water pressure varies greatly with increasing depth. At the surface of water at sea level the weight of air above the water is considered as one atmosphere which is equivalent to pressure exerted by a 10 m high column of sea water. At a depth of 9,750 m, one of the greatest depths known in the Pacific, the pressure is 962 atmospheres. Yet, this tremendous weight does not crush bottom organisms living there. The limiting factor for life at great depths is not pressure but paucity of food.

Also waves that break as surf on the beach have a pounding force which can reach high values. The average force of such waves of the North Sea has been reckoned at 1.5 kg per sq. cm or 15,000 kg per sq. m. This means that living organisms inhabiting rocky, coastal waters must be equipped with protective covering or have a special ability to avoid being crushed by such forces. The force of the surf is less on sandy beaches than on rocky coasts. Wave motion does not only make the surface turbulent,

in coastal waters it is detectable down to about 40 m which is the case on the Algerian coast, but in other areas near land the lower extent goes down to about 200 m. However, in the open ocean the wave motion extends to depths of several hundred metres.

The oscillating movements of tides, caused by cosmic forces, are another important feature of the oceanic environment along the coasts. They contribute to the chemical mixing of the water and extend to great depths. Near the Canary Islands, for example, this depth extends to 1,800–2,000 m.

Biotic divisions and zonation

Ecologically the ocean offers a wide range of habitats with varied chemical, physical and biotic factors. Life zones exist in practically all parts of the sea, from deep bottoms to the surface. Rocks, stones, gravel, sand, mud, shells and corals constitute different environments, each with a distinctive fauna varying with depths, salinity, tidal rhythms, and so on. Most marine organisms float or swim in the pelagic zone, where the greatest abundance of life forms is found.

In coastal areas of tropical seas environmental conditions for marine organisms are optimal and the diversity in species is greatest.

Marine biotic divisions of the ocean are the pelagic (pelagial) and benthic (benthal) life zones. Pelagic refers to a marine region consisting of free water layers of

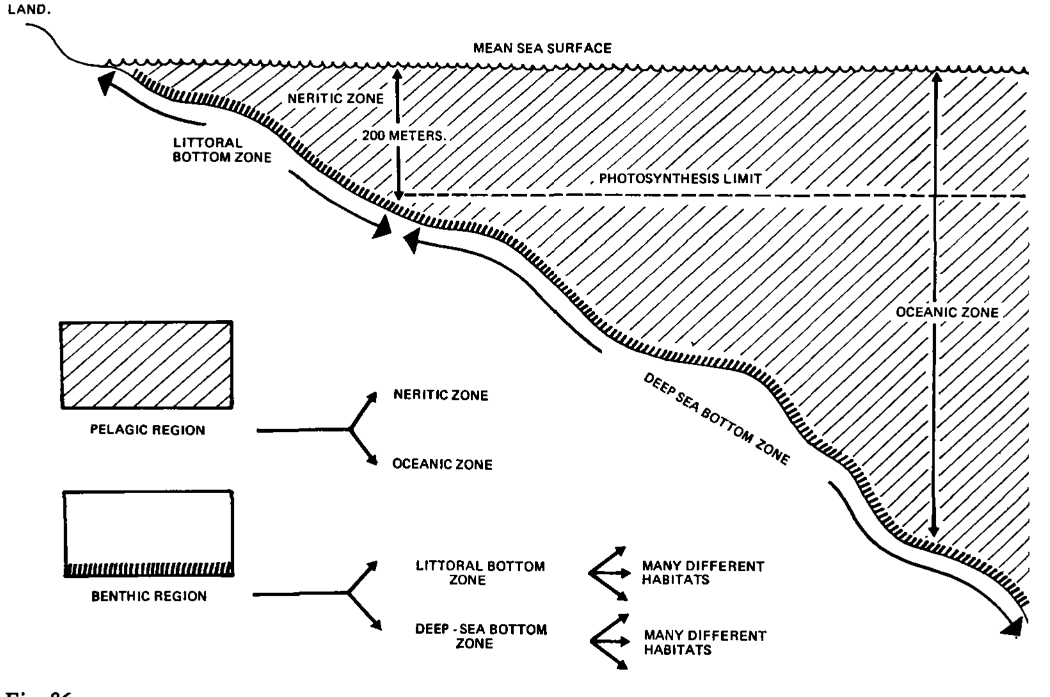


Fig. 86
Oceanic life zones

the open sea, while benthic is the bottom region, where animals and plants find physical support other than just the water. Both regions fall vertically from the shoreline to the greatest depths, which means that the pelagic region increases to a tremendous volume, particularly outside the continental shelf. The latter lies at an average depth of about 200 m, but in many coastal areas there is no shelf at all, resulting in a dramatic contrast between the rocky shoreline and the great depths immediately below.

The pelagic and benthic regions merit the term 'marine biomes'. The former may be divided, despite its uniformity, into several habitats distinguished by different degrees of light penetration from the surface, currents and other factors. Each of these habitats are characterized by certain animals, but many of them move vertically up and down within a rhythm of twenty-four hours. We shall not go into these distinctions here, but there is a marked environmental difference in both the pelagic and benthic regions at a depth of about 200 m which sets the limit for the photosynthesis. In the pelagic region the area above the shallow littoral zone is called the neritic zone while the area of the pelagic region above the benthic region is called the oceanic zone (Fig. 86).

The littoral habitats are of course much influenced by the tides, which regularly shift them between two elements—the hydrosphere and the atmosphere—and two ecosystems, one marine and the other terrestrial. Nevertheless several organisms living in the tidal zone are sedentary and have adapted to the highly variable conditions.

There is also a 'spray zone' above the high tide mark, reached by saltwater only through waves and spray, but it is enough to support marine air-breathing animals which are protected from desiccation by a shell.

In a vertical sense the benthic region shows a distinct zonation within each habitat: the biocommunities of the solid bottoms (rocks, stones, pebbles, corals) contrast greatly to those of loose bottoms (mud, sand, pulverized shells) at the same depth. To these structural differences of the benthic region come various types of plant growths in the littoral zone. Algal vegetation begins near the low tide level on both rocky and sandy bottoms. Some marine algae grow to gigantic size reaching the surface from depths down to 80 m (off South America), virtually forming a 'sea forest'. There are also marine meadows consisting of dense stands of various genera belonging to five families, *viz. Ruppiaceae*, *Zosteraceae*, *Posidoniaceae*, *Zannichelliaceae* and *Hydrocharitaceae*. The tropical-subtropical species are quite distinct from those of the temperate and boreal regions. Of the Indo-West Pacific species, *Enhalus acoroides*, *Halodule pinifolia*, *H. tridentata*, *Thalassia hemprichii* have wide distribution from East Africa to the Indo-Malay region and the Pacific islands while *Halophila stipulacea* is confined to the Red Sea and West Indian Ocean. The temperate-boreal group is characterized by the genera *Althenia*, *Posidonia* and *Zostera*. Of these *Zostera* has the most extensive range with distribution on both east and west coasts of North America, on the coasts of Atlantic Europe, east Asia, Australia and South Africa.

All these marine plants add significantly to the ecological setting of the various littoral habitats. Many animals are restricted to these plant beds, and others use them for spawning, depositories for eggs and nurseries for larvae. The result is that rich

animal communities develop in seaweed thickets where they find abundant oxygen, plenty of food, shelter and body support without resting on the bottom. The abyssal zone is less varied than the littoral zone. Mud bottoms dominate, nourished by sedimentation during vast periods of time.

In the pelagic region as a life zone the most characteristic feature is that the organisms living there are able to maintain themselves in the open water without floating up to the surface or sinking to the bottom. Another characteristic of the pelagic region is the passively floating or suspended plankton, although many of these microscopic animals as well as the free-swimming nekton are able to rise to the surface each night.

The two pelagic regions have partly different faunas. The more diversified neritic zone is inhabited by many species which are dependent on the bottom during at least part of their life cycle, while the more uniform oceanic zone is distinguished by animal species which are entirely independent of the benthic region. In the deep-sea world of the oceanic zone there are no plants other than occasional debris dropping from upper layers. This means that there are no herbivores below a depth of about 200–300 m. All animals there are either predators or scavengers.

The absence of cosmic light in the deep sea is partly neutralized by the production of light by a relatively high number of animals. No less than 44 per cent of fish living below a depth of 900 m are light producers and about two-thirds of the bathypelagic fish species are luminous. The abyssal zone virtually swarms with animals producing light. This means that bioluminescence has an important biological function. Is it for the attraction of prey, or warning signals for predators or does it function as food lures or sexual recognition?

Food chains and productivity

The productivity of the continental shelves is many times higher than that of the great depths outside them. This means that most parts of the oceans are not so productive that they can be considered as immense food reserves of tomorrow for the starving human population of the world.

Tropical marine communities live in a stable environment with high temperatures. Reproduction is continuous and takes place at a high rate with most species breeding throughout the year. Some species have one long breeding season while others two. Despite the fact that tropical seas are rich in species with a high rate of energy flow, they have in surface waters mostly a smaller number of individuals per species than is the case in many temperate and polar seas. In temperate regions the surface layers are well provided with nutrients by convection during the winter and early spring but as the supply of nutrients diminishes towards summer, the fertility also declines. Fertile areas also result by up-welling in the currents along the eastern part of the Atlantic, Indian and Pacific oceans at low latitudes. Throughout the tropics where a permanent thermocline is present and the mixing is minimal, productivity is low despite rapid regeneration, e.g. the Sargasso Sea.

In terms of organic reproduction, Steeman Nielsen and Aabye recognize four different regions; those with a very considerable admixture of nutrient-rich water to the photosynthetic zone; the organic production is high amounting to about 0.5–3.0 $gC/m^2/day$, e.g. southern part of the Benguela current; those with a fairly steady admixture of nutrients giving production value of 0.2–0.5 $gC/m^2/day$, e.g. areas of divergence caused by equatorial counter-currents in the Atlantic, Indian and Pacific oceans; those with small influx of water from medium depths into the euphotic zone by turbulence; the production is 0.1–0.2 $gC/m^2/day$, e.g. most areas in the tropics and subtropics; those with old surface water without regular influx or renewal. Production here is about 0.05 $gC/m^2/day$, e.g. the Sargasso Sea.

In the sea are many food chains, the links consisting of vast varieties of plankton, mollusks, arthropods, fish, birds and marine reptiles and mammals. Plankton is the mass of small animal and plant life that drifts with the currents and tides of the sea—small animals such as protozoans, larval fish, and crustaceans, some larger ones such as jellyfish, and tiny plants such as algae. Plankton organisms dominate the sea populations in number and form the base for almost all the food chains. In addition, many fish and mammals utilize plankton directly. For example, two sharks that are the largest fish in the world and the baleen whales, the largest mammals that have ever existed, feed directly on plankton. Man likewise utilizes plankton as food, as transferred through other groups of marine animals. However, the energy loss through the different links of food chains is considerable. The marine food chain often contains more steps than the terrestrial one as shown in Table 44.

Table 44
Links of food chains

Trophic level	Marine food chain	Terrestrial food chain
Producers	Phytoplankton	Grasses, leaves, twigs
Primary consumers (1) (herbivores)	Zooplankton	Insects, rodents, antelopes
Secondary consumers (2) (predators)	Sardines and other small carnivorous fish	Predatory birds, carnivores, man, and others
Tertiary consumers (3) (predators)	Mackerels and other intermediate carnivorous fish	
Quarterly consumers (4) (predators)	Tunas, bonitos, sharks and other top carnivorous fish	

The most productive areas of the ocean are those where nutritive mineral elements give optimal growth conditions to planktonic plants. However, we still do not know enough about ocean productivity. Since evaluations and estimates have been based on fishery and whaling catches, they relate only to particular species, excluding the majority of marine life forms. The seas are vast pastures, but the grazing efficacy of the

vertebrates on phytoplankton (plants) and zooplankton (animals) is little known. Moreover, the distribution of zooplankton is patchy, and their quality varies greatly. In the sea only a few species of economically important fish eat plants directly. Copepods form the main link in the food chain between marine plants and fish. They are small planktonic crustaceans occurring in all the oceans and many of them graze phytoplankton. In coastal waters planktonic stages of bottom-living worms, echinoderms such as starfish, and molluscs feed on phytoplankton. Thus zooplankton grazing on phytoplankton have a key position in the food chain of economically important marine species.

In comparing the primary productivity of the oceans with that of the land and freshwater, the highest production of organic carbon per year, expressed in metric tons per sq. km, is found in the sea, namely the Atlantic Ocean with 1,000 tons, while a corn field in Ohio reached 862 tons and a lake in Wisconsin 480 tons (Odum, 1971). Yet man's harvest of the secondary productivity of the sea is low in comparison to that of the land. The latter is consumed in a more direct way at steps 1 and 2, while the marine harvest is taken at steps 2–4 (Table 44). Man is not like the larger whales able to feed directly on plankton.

Marine resources utilized by man are entirely self-sustaining. They require neither fertilizers nor cultivation and have a high reproductive rate. Therefore, management of marine ecosystems should be aimed at maintaining them intact and harvesting fish, turtles and invertebrates at maximum size and on sustained yield basis. However, some species may have higher commercial value at an inferior size or in the form of eggs. Such harvests will reduce the biomass.

The importance of adequate food densities emphasizes the vulnerability of marine ecosystems and their productivity. If just one link in the food chain of a life cycle within a region is damaged, the whole community will suffer. The population sizes of such important fish as herrings, sardines, and anchovies may be influenced negatively for decades. Human exploitation of ocean resources must therefore be intelligent, based on a full knowledge of their productivity, life cycles, population shifts, and migrations (both vertical and horizontal), so that the annual harvest is synchronized with all these phenomena without diminishing the capital by reducing the breeding rate—in other words, so that sustained yield will be obtained. This means, for example, that one or several countries should not exploit in different seasons of the year the same migratory populations of marine animals in two or more regions without being aware that in reality only one stock is involved.

Steady population decline of a marine species usually reflects over-exploitation. Decrease in abundance often leads to more intensive fishing methods, for instance, going to finer-meshed trawls that catch smaller sizes of the declining species. The proportion of immature fish in the catch increases, with a direct negative effect on the reproduction of the species. The population dwindles and so does the mean size of the fish caught. On the other hand, reductions of local populations through fishing may have a positive effect on the mean size of the species involved because the intra-specific competition decreases and more food becomes available for each individual.

Overfishing around Asia has occurred and is still going on in many sea regions. This is the case in South China Sea and the straits of Singapore and Malacca.

Marine reptiles are also part of the precious resources of oceans. Sea turtles spend most of their lives in the open sea, though some species are confined to shallow coastal waters. However, all species have to reproduce on land, and therefore all of them occur, at least temporarily, in coastal waters. The necessity for marine turtles to go ashore during the breeding season to lay eggs makes the adult females very vulnerable, and that is also the case with the eggs and the newly hatched young. In addition, exploitation of shore areas and alterations in shorelines that serve as breeding sites for sea turtles reduce the possibilities of reproduction. The green turtle (*Chelonia mydas*) is herbivorous, grazing on manatee grass and turtle grass. It furnishes meat of unsurpassed quality and could, if managed properly, constitute a very important food resource in tropical and subtropical countries. Its protein supply potential has been underestimated, and, in most countries where it is available, almost entirely neglected by governments, despite the fact that for centuries the sea turtle has been a considerable part of the diet of human coastal populations. For centuries too a careless exploitation of sea turtles has been taking place, with the result that several species have been severely depleted in many areas where formerly they had been abundant. This is especially the case around Malaysian shores where, today, far fewer turtles come for breeding.

The first step would be to restore the populations to the optimal level that the environment can produce. Artificial hatching and raising are possible and have already been undertaken successfully, and regular releases of young green turtles into the sea are now made in many parts of the world. But this work is only a start.

Disastrous effects on marine life zones are continuously caused by pollution, such as industrial effluents and oil discharges from ships in the open sea. Like the extermination of whales, this results from the unfortunate situation that international waters are considered to be a 'no man's sea', where neither nations nor individuals have responsibility.

In 1971, the French explorer J. Y. Cousteau made at the Council of Europe (celebrating the European Conservation Year) a statement on global marine degradation, based on his observations of over nearly thirty years, extending to depths of at least 500 m in a wide range of oceans and seas, and at the conclusion of a voyage of 248,000 km. Cousteau draws his experience from repeated observations in the Mediterranean, the Red Sea, the Indian Ocean, the southern Atlantic, the Caribbean Sea and the Pacific Ocean. The essence of his message is that the intensity of life in the oceans and seas has in thirty years diminished by more than 30 per cent, probably by an average of 40 per cent, and possibly by nearly 50 per cent in the past twenty years. The main causes are pollution, over-fishing and environmental changes resulting from human intervention.

Estuaries, mangroves and coral reefs

Tidal estuaries and brackish water habitats produce special environmental conditions that vary greatly from one area to another, depending on tides and freshwater outflows, temperature and salinity, substratum, vegetation, etc.

In the tropics mangrove growths represent a very special environment, particularly below the water surface where the tangle of mangrove roots creates a peculiar underwater world with an extraordinary gathering of marine and freshwater animals. Usually estuaries and mangrove forests are regarded as unproductive wastelands, suitable for dumping industrial and urban wastes. On the contrary, estuaries are often among the world's most fertile areas.

A basic factor underlying the productivity of an estuarine ecosystem is the tidal action in nutrient cycling. The natural productivity is of such magnitude that management emphasis has to be on utilization rather than production. The fertility and productivity of investigated estuaries exceed ten times the corresponding values for ecosystems such as grasslands, forests, eutrophic lakes and ordinary agriculture. The data obtained by Odum in Georgia are believed to be applicable throughout the Atlantic and Gulf coasts of temperate and subtropical North America. If this is true, estuaries in other areas of the world should likewise be highly productive. Odum estimates that the Sapelo estuarine marshes in Georgia have a gross primary net production of about 2,000 grammes of dry matter per square metre per year. This is twenty tons per ha. The highest yields of wheat and corn obtained in northern Europe are about fourteen tons of dry matter per ha. World average net production of wheat is about three tons of dry matter per ha per year. In terms of potential food energy usable by organisms, the net production of the Georgian estuaries investigated by Odum is apparently in the order of 8,000 calories per square metre, or about 80 million calories per ha per year. To emphasize its significance in relation to the food requirements of man it may be mentioned that the mean quantity of food to keep a human being healthy is estimated to be 2,750 kc a day on a world population basis.

Several major mechanisms explain why estuaries are so fertile and productive. The meeting of waters of different salinities produces a vertical mixing of nutrients. Instead of being swept out, the nutrients move up and down, cycling rapidly among organisms, water and bottom sediments. Secondly the back and forth tidal flow continually supplies nutrients and oxygen and automatically removes waste products, so that organisms need not use their own energy for these purposes. Perhaps the most important finding on estuaries, next to how extraordinarily productive they are, is that the entire estuary marsh complex—sounds, creeks, mud and sand flats—must be considered as a unit, regardless of whether it is one or 100 miles wide. The prime importance to humans of the tremendous productivity of tidal estuaries seems to be as a source of sea food. Estuarine wetlands continuously fertilize the pastures and bottoms of the sea and are in this way a basic link in the marine food chain. It is certainly more useful for man to leave potentially useful tidal marshes and estuaries intact, instead of converting them into less productive lands for agriculture and buildings.

Apparently estuaries and brackish waterways well protected from the open sea

are a natural resource that is entirely neglected or badly used in most places where they exist. Brackish inshore water may even be farmed as in Indonesia and Singapore. The potential value of such agriculture may be judged by comparing the productivity of already existing commercial brackish ponds and conventional farmlands. In tropical Asia such ponds yield from 160 to 1,500 kg of fish and prawns per ha per year. The annual total edible protein, expressed in dry weight, is up to 81 kg per ha in brackish ponds, while the grain produced per ha under average agricultural conditions in the United States will feed enough livestock to yield about 21 kg of protein food a year.

One must bear in mind that modern agriculture is based upon intensive research into all its aspects, while aquaculture has virtually no formal study underlying its methods. Yet there are ponds in South-East Asia where, in addition to the rice crops, 750 kg of prawns, net weight, are harvested annually. In such cases the rate of brackish-water farm yield is much higher than that of agriculture and even a better fishing rate than that obtained by commercial fishermen. Undoubtedly the protein harvest from brackish waters can be multiplied by management measures that are much less costly than what is required to maintain the productivity of conventional farmlands.

Mangroves occur along marine shores of tropical and subtropical seas but also grow in deltas, estuaries, brackish-water swamps and up rivers as far as the tide carries salt-water. They are best developed in the Indo-Pacific region with the most luxuriant growth in the Indo-Malayan archipelago. They are, in fact, land plants which have evolved a high tolerance to the salinities of the open sea.

Mangroves are important land and island builders which not only extend coasts but also protect them from erosion and hurricanes. Moreover, as indicated earlier, mangroves provide habitats for numerous marine organisms and constitute feeding, breeding and nursing grounds for many species of fish, molluscs and crustaceans. The nutrients released by the mangrove litter are the basis for microscopic diatoms and blue-green algae which, in turn, are the food for many organisms, both invertebrates and fish. These are the first links of the food chain and, therefore, essential for organisms higher up the same food chain. Terrestrial vertebrates such as mammals, birds and reptiles also utilize mangrove habitats. In fact, mangrove swamps comprise habitats which belong to the most productive in the world. They are in many areas of importance to commercial and sport fishing. Some mangroves provide a source of tannin which is used for making leather, ropes, nets and sails. They are also used for lumbering, domestic fuel and charcoal. Despite the growing appreciation of mangrove values, large areas of such forests have been, and still are, destroyed in many parts of the world.

In a biological sense coral reefs are a world asset. They are important economically because the productivity of their organisms is high and contributes to several food chains in the ocean. Their fishery value is also high, both directly and indirectly, for they provide food, shelter, and spawning sites and nurseries for marine life. In addition they have mineral and recreational values. Coral reefs often form enormous natural breakwaters extending from sea bottom to the surface layer and covering large areas as a result of the work of tiny animals through long periods of time. Corals are colonial animals which constantly secrete lime to build reefs. The subtropical or tropical

climate is a condition for the existence of living coral reefs. Some coral reefs are estimated to be some 30 million years old. Others consist of about 300 species of coral in all various shapes and sizes. Coral colonies constantly increase the number of living skeletons and in this way maintain the reefs in balance against the forces which break them down. Corals usually feed on zooplankton at night.

The productivity of coral reefs is directly linked to an undisturbed richness and diversity of the fish and invertebrate fauna, because tropical waters are low in the concentration of plant nutrients containing nitrogen and phosphorus. Therefore, the high productivity of a coral reef is sustained by a constant recycling of nutrients produced by a varied and complex biocommunity. Heavy fishing interferes with the cyclic return of nutrients, because it removes a vital part of the ecosystem. The algae production decreases. Corals live in symbiosis with unicellular algae and are particularly sensitive to habitat disruption. The loss of nutrients in coral reefs cannot be replaced by nutrients from the open sea. The situation is somewhat better in coral reefs surrounding totally enclosed lagoons that retain water at low tide.

As structures, coral reefs are vulnerable to human destruction. Prospecting for offshore petroleum deposits may lead to oil-drilling activity in those waters, with consequent risks of spillage that would be highly dangerous to marine life in general and to coralreef habitats in particular.

Conservation efforts and management planning of coral reefs, based on ecological research, would in the long run benefit mankind in the form of dividends from a chain of marine renewable resources.

10 Food and nutrition

Nutritive content of different foods

The nutritionist usually describes the chemical composition of food in terms of its relative content of various proximate food substance. A list of constituents of some common food items is given in Table 45.

Dietary needs

A diet may be defined as the kinds of food on which a person or group lives. A diet must include adequate amounts of proteins, carbohydrates, fats, water, vitamins and minerals for proper nutrition of the body. A diet which supplies all these essential nutrients in correct proportions and in adequate quantities is known as a 'balanced diet'.

Dietary needs of an individual depend upon age, sex, size, level of activity, health and climate. Thus, it is difficult to establish a standard diet for all people. However, FAO has established standards, defined as safe practical allowances for calories, proteins and calcium, intended for international use. Calorie allowances are based on the FAO-defined reference man and woman. The reference man and woman are both 25 years of age, weigh 65 and 55 kg respectively, are healthy and moderately active. They live in an environment with a mean annual temperature of 10° C, consume an adequate well-balanced diet and neither gain nor lose weight. Equations for determining energy requirements (in calories) were developed based on the criteria that 75 per cent of the energy consumption is directly related to the body weight and the remaining 25 per cent are independent of body weight as given below:

$$\text{Men: } E = 815 + 36.6\ W$$
$$\text{Women: } E = 580 + 31.1\ W$$

where E is energy requirement in calories and W is weight in kg.

Provision is made also for age and temperature adjustments. A deduction of 3 per cent of the calorie requirement is made for every decade from age 20 through 50;

Table 45. Relative nutritive constituents of common foods

Name	Edible portion %	Per 100 grammes of edible portion										
		Protein g	Fat g	Carbo-hydrate g	Vit. A U.I.	Vit. B_1 mg	Vit. B_2 mg	Nicotinic acid mg	Vit. C mg	Ca mg	Fe mg	kc
Cereals												
Rice raw, milled	100	6.8	0.5	78.2	0[1]	0.09	0.03	1.9	0	10	3.1	345
Wheat flour	100	12.1	1.7	69.4	49	0.49	0.29	4.3	0	48	11.5	341
Corn (maize)	—[2]	—	—	—	—	—	—	—	—	—	—	—
Milo (Juar)	100	10.4	1.9	72.6	79	0.37	0.28	1.8	0	25	5.8	349
Bajra	84	11.6	5.0	67.5	220	0.33	0.16	3.2	0	42	14.3	356
Pulses and legumes												
Bengalgram (dhal)	100	20.8	5.6	59.8	216	0.48	0.18	2.4	1	56	9.1	372
Black gram (Udad dhal)	100	24.0	1.4	59.6	64	0.42	0.37	2.0	0	154	9.1	347
Green gram (Mung dhal)	100	24.5	1.2	59.9	83	0.72	0.15	2.4	0	75	8.5	351
Lentil (Masur Dhal)	100	25.1	0.7	59.0	450	0.45	0.49	1.5	0	69	4.8	343
Red gram (Arhar)	100	22.3	1.7	57.6	220	0.45	0.51	2.6	0	73	5.8	355
Peas, dry	100	19.7	1.1	56.5	66	0.47	0.38	1.9	0	75	5.1	315
Beans, dry	—	—	—	—	—	—	—	—	—	—	—	—
Nuts and oil seeds												
Gingelly seeds (Til)	100	18.3	43.3	25.0	100	1.01	0.06	4.4	0	1 450	10.5	563
Ground-nuts	—	26.7	40.1	20.3	63	0.90	0.30	14.1	0	50	1.6	549
Roots and tubers and bulbs												
Beet root	85	1.7	0.1	8.8	0	0.04	0.09	0.4	88	200	1.0	43
Carrot	95	0.9	0.2	10.6	3 150	0.04	0.02	0.6	3	80	2.2	47
Onion, big	—	1.2	0.1	11.0	0	0.08	0.01	0.4	11	180	0.7	49
Potato	100	1.6	0.1	22.6	40	0.10	0.01	1.2	17	10	0.7	97
Radish, white	99	0.7	0.1	3.4	5	0.06	0.02	0.5	15	50	0.4	17
Sweet potato	—	1.2	0.3	28.2	10	0.08	0.04	0.7	24	20	0.8	120
Yam, elephant (Zamin Kand)	—	1.2	0.1	18.4	434	0.06	0.07	0.7	0	50	0.8	79
Leafy vegetables												
Bathua leaves (Chandan bathwa)	—	3.7	0.4	2.9	4 680	0.01	0.12	0.6	32	150	6.0	30
Cabbage	51	5.1	0.5	13.1	2 000	0.06	0.03	0.4	124	39	0.8	77
Colocasia	—	3.9	1.5	6.8	23 000	0.07	0.23	1.2	12	290	14.3	56
Coriander leaves (dhania)	70	3.3	0.6	7.5	11 530	0.05	0.06	0.8	135	184	18.5	48
Fenugreek (Methi)	59	4.4	0.9	6.0	6 450	0.05	0.15	0.7	54	360	17.2	49
Gogu (Patwa)	76	1.7	1.1	9.9	4 830	0.07	0.21	1.5	20	172	5.0	56
Mint (Paudina)	45	4.8	0.6	5.8	2 700	0.05	0.08	0.4	27	200	15.6	47
Spinach (Palak)	87	2.0	0.7	2.9	9 300	0.03	0.07	0.5	28	73	10.9	26
Other vegetables												
Brinjal	91	1.4	0.3	4.0	124	0.04	0.11	0.9	12	18	0.9	24
Cauliflower	70	2.6	0.4	4.0	51	0.04	0.10	1.0	26	33	1.5	30

Double beans	—	8.3	0.3	12.3	—	—	—	—	22	40	2.3	85
Drumstick	83	2.5	0.1	3.7	184	0.05	0.07	0.2	120	30	5.3	26
Kovai fruit (kundree)	96	1.2	0.1	3.1	260	0.07	0.08	0.6	22	40	1.4	18
Lady's fingers	84	1.9	0.2	6.4	88	0.07	0.10	0.6	13	66	1.5	35
Pumpkin	79	1.4	0.1	4.6	84	0.06	0.04	0.5	2	10	0.7	25
Green tomato	98	1.9	0.1	4.3	320	0.07	0.01	0.4	31	20	1.8	20
Fleshy fruits												
Apple	90	0.3	0.1	13.3	0	0.12	0.03	0.2	2	9	1.0	55
Banana, green variety, peeled	—	0.8	0.8	24.4	90	0.03	0.03	0.5	0	26	2.1	107
Figs (Angeer)	99	1.3	0.2	7.6	270	0.06	0.05	0.6	5	60	1.2	37
Grapes, pale green	—	0.5	0.3	16.5	0	—	—	0	1	20	1.5	71
Guava	100	0.9	0.3	11.2	0	0.03	0.03	0.3	212	50	1.2	51
Jack fruit	30	1.9	0.1	19.8	292	0.03	0.13	0.4	7	20	0.5	88
Lemon, sour	97	0.6	0.7	8.2	0	—	—	0.1	26	100	2.4	42
Jambu	75	0.7	0.3	14.0	80	0.03	0.01	0.2	18	15	1.2	62
Sweet lime (musammi)	71	0.8	0.3	9.3	0	—	—	0	50	40	0.6	45
Mango, alphanso	—	0.4	0.9	16.3	15 220	0.08	0.09	4.1	25	42	1.6	74
Water melon	78	0.2	0.2	3.5	0	0.02	0.04	0.1	1	11	7.9	16
Orange (Nagpur)	70	0.6	0.2	8.9	1 800	—	—	0	30	20	0.5	40
Papaya, ripe	75	0.6	0.1	7.2	1 110	0.04	0.25	0.2	57	17	0.5	32
Pineapple	60	0.4	0.1	10.8	30	0.20	0.12	0.1	39	20	1.2	46
Tomato, ripe	100	0.9	0.2	3.6	585	0.12	0.06	0.4	27	48	0.4	20
Animal foods		0										
Fish, Bugda	—	18.8	1.6	4.5	—	—	—	—	—	290	1.4	108
Crab, small	—	11.2	9.8	9.2	—	—	—	—	—	1 606	21.2	170
Lobster	—	20.5	0.9	—	—	—	—	—	—	16	—	90
Pomfrets, white	68	17.0	1.3	1.8	—	—	0.55	2.6	—	178	12.2	87
Prawn	—	21.5	1.7	1.2	—	0.09	0.10	4.8	—	38	—	101
Beef	—	22.6	2.6	—	60	0.15	0.04	6.4	2	10	0.8	114
Egg-hen	—	13.3	13.3	—	2 200	0.10	0.18	0.1	0	60	2.1	175
Goat meat	—	21.4	3.6	—	—	—	—	—	—	12	—	118
Liver, sheep	—	19.3	7.5	1.4	22 300	0.36	0.70	17.6	20	10	6.3	150
Mutton	—	18.5	13.3	—	—	—	—	—	—	150	2.5	194
Pork	—	18.7	4.4	—	0	0.54	0.09	2.8	2	30	2.3	114
Milk, cow	100	3.2	5.1	4.4	165	0.05	0.18	0.1	2	149	0.2	67
Milk, buffalo	100	4.3	8.8	5.1	160	0.04	0.10	0.1	3	210	0.2	118
Curds	100	3.1	4.0	2.9	102	0.05	0.16	0.1	1	149	0.3	60

1. 0=absent.
2. —=not estimated.

Source: S. C. Bala Subramaniam (Nutrition Research Laboratories, Indian Council of Medical Research, Hyderabad). *Nutritive Value of Common Food Stuffs, You and Your Health*, New Delhi, National Book Trust, 1969.

7.5 per cent for each decade between 50 and 70 and 10 per cent above 70 are recommended. For temperature adjustment, a 5 per cent deduction for every 10° C above reference temperature and a 3 per cent addition for every 10° C below reference temperature are suggested.

Protein allowances, as recommended by FAO, are based on three sets of criteria: average requirement of the population; average requirement plus 20 per cent as a safety factor, and average requirement minus 20 per cent. Table 46 provides data on protein requirements for various age groups.

The figures in Table 46 are based on a NPU (net protein utilization) of 100, so the amounts given should be adjusted for the NPU of diet usually eaten by a given population, where NPU=N retained/N intake.

Table 46
Protein requirements for various age groups

Age	Protein requirements FAO reference protein: g/kg		
	I[1]	II	III
Infants (months)			
0–3			2.3
3–6			1.8
6–9			1.5
9–12			1.2
Children (years)			
1–3	0.70	0.88	1.06
4–6	0.65	0.81	0.97
7–9	0.62	0.77	0.92
10–12	0.58	0.72	0.86
Adolescents			
13–15	0.56	0.72	0.84
16–19	0.51	0.64	0.77
Adults	0.41	0.59	0.71

1. I=Inadequate for all but a very small fraction (2.5 per cent of the population).
II=Average protein requirements of the population.
III=Adequate for all but a very small fraction (2.5 per cent of the population).

Source: FAO protein requirements (1965).
Pike and Brown. *Nutrition: an Integrated Approach*. New Delhi, Wiley Eastern (Private) Ltd, 1970, p. 463.

Table 47
Practical allowance for calcium (FAO)

Age	mg/day
0–12 months[1]	500–600
1–9 years	400–500
10–15 years	600–700
16–19 years	500–600
Adults	400–500

1. Artificially fed.

Source: FAO protein requirements (1965).
Pike and Brown. *Nutrition: an Integrated Approach*. New Delhi, Wiley Eastern (Private) Ltd, 1970, p. 464.

A calcium allowance, compatible with health and attainable in many areas of the world, was also suggested by FAO. This is given in Table 47. For pregnant and lactating women, allowances of 1,000 and 2,000 mg respectively are recommended.

Recommended daily dietary allowances of nutrients for Indians are shown in Table 48. As the Indian population contains both vegetarian and non-vegetarian groups, the nutrition expert group of the Indian Council of Medical Research has recommended both types of balanced diets for different age groups as shown in Table 49.

Dietary patterns

The dietary pattern varies widely in different parts of the world. It is generally developed around the kinds of food produced depending upon the climatic conditions of the region, economic capacity, religions, customs, taboos, tastes and habits of the people.

The type of cereal consumed as staple food varies. Asian people are mostly rice-eaters. They produce and consume about 90 per cent of the world's rice.

The types of animal food in diets likewise depend upon locally available animal products.

The consumption level for different foods in the types of diet commonly found in Asia and the Far East has been recently summarized by the Indicative World Plan for Agricultural Development (Provisional Regional Study No. 4, Asia and the Far East, Vol. I, FAO, Rome, 1968) as follows:

I. *Most of India, Bangladesh and Sri Lanka*	
Rice, millets and other cereals	Moderately high
Pulses, fats and oils	Moderate
Milk	Low
Meat, fish and eggs	Very low
II. *Thailand, west Malaysia, Philippines and the Republic of Korea*	
Rice, starchy roots and tubers	High
Meat, fish and eggs	Moderate
Milk and pulses	Low
III. *India (Punjab) and Pakistan*	
Wheat, rice	High
Milk and pulses	Moderate
Meat, fish and eggs	Low

The predominant features of the diet in the region, as summarized by ECAFE's Asian Population Studies Series No. 11, are the following: cereals contribute up to 70–80 per cent of both the calorie and the protein supplies; pulses are a significant source of protein in southern Asia, contributing about 20 per cent of protein and about 10 per cent of calories; foods of animal origin contribute only 4 per cent of calories and 10 per cent of proteins in southern Asia; however, their contribution in eastern Asia is higher at 8 per cent of calories and 28 per cent of proteins; among foods of animal origin, milk is more significant in southern Asia while meat and fish are more important in the rest of the region.

Table 48. Dietary allowances of nutrients for Indians[1]

Person	Weight Age	Level of activity	Net kc	Protein g	Calcium g	Iron mg	Vitamin A U.I.	Thiamine mg	Ascorbic acid mg	Vitamin D U.I.
Man	55 kg	Sedentary work	2 400	55	1.0					
	55 kg	Moderate work	2 800	55	1.0					
	55 kg	Heavy work	3 900	55	1.0					
Woman	45 kg	Sedentary work	2 000	45	1.0	20	3 000	1.0		400
	45 kg	Moderate work	2 300	45	1.0	to	to	to	50	to
	45 kg	Heavy work	3 000	45	1.0	30	4 000	2.0		800
		Pregnancy (latter half)	2 300	100	1.5					
		Lactation	2 700	100	2.0					
Children	0–6 months		120×kg	3.5×kg						
	7–12 months		100×kg	3.5×kg						
	1–3 years		1 200	3.5×kg						
	4–5 years		1 500	3.5×kg						
	5–6 years		1 500	3.0×kg						
	6–7 years		1 800	3.0×kg	1.0	10	3 000	0.5	30	400
	7–9 years		1 800	3.0×kg	to	to	to	to	to	to
	10–12 years		2 100	3.0×kg	1.5	30	4 000	1.0	50	800
Adolescents	13–15 years	Girls	2 100	3.0×kg						
		Boys	2 500	3.0×kg						
	16–19 years	Girls	2 100	2.0×kg						
		Boys	3 150	2.0×kg						

1. Recommended by the nutrition advisory committee of the Indian Council of Medical Research, 1944; revised in 1958.

Source: Gopalan and Narasinghrao. *Dietary Allowances for Indians.* Indian Council of Medical Research, Special Report Series, No. 60. Nutrition Research Laboratory, Hyderabad-7.

Table 49. Balanced diet for different age groups of Indian population[1]

	Age group	Vegetarian or non-vegetarian	Cereals g	Pulses g	Green leafy vegetables g	Other vegetables g	Roots and tubers g	Fruits g	Milk g	Fats and oils g	Meat and fish g	Eggs g	Sugar g	Ground-nuts[2] g
Adult male	Sedentary work	VEG	400	70	100	75	75	30	200	30	—	—	30	—
		NON	400	55	100	75	75	30	100	40	30	30	30	—
	Moderate	VEG	475	80	125	75	100	30	200	40	—	—	40	—
		NON	475	65	125	75	100	30	200	40	30	30	40	—
	Heavy work	VEG	650	80	125	100	100	30	200	50	—	—	55	50
		NON	650	65	125	100	100	30	100	50	30	30	55	50
Adult female	Sedentary work	VEG	300	60	125	75	50	30	200	30	—	—	30	—
		NON	300	45	125	75	50	30	100	35	30	30	30	—
	Moderate	VEG	350	70	125	75	75	30	200	35	—	—	30	—
		NON	350	55	125	75	75	30	100	40	30	30	30	—
	Heavy work	VEG	475	70	125	100	100	30	200	40	—	—	40	—
		NON	475	55	125	100	100	30	100	45	30	30	40	—
	Pregnancy		50	—	+25	—	—	—	+125	—	—	—	+10	40
	Lactation		100	+10	25	—	—	—	+125	+15	—	—	+20	40
Children	1–3 years	VEG	150	50	50	30		50	300	20	—		30	—
		NON	150	40	50	30		50	200	20	30		30	—
	4–6 years	VEG	200	60	75	50		50	250	25	—		40	—
		NON	200	50	75	50		50	200	25	30		40	—
	7–9 years	VEG	250	70	75	50		50	250	30	—		50	—
		NON	250	60	75	50		50	200	30	30		50	—
	10–12 years	VEG	320	70	100	75		50	250	35	—		50	—
		NON	320	60	100	75		50	200	35	30		50	—
Adolescent boys	13–15 years	VEG	430	70	100	75	75	30	250	35	—	—	30	—
		NON	430	50	100	75	75	30	150	40	30	30	30	—
	16–18 years	VEG	450	70	100	75	100	30	250	45	—	—	40	50
		NON	450	50	100	75	100	30	150	50	30	30	40	50
Adolescent girls	13–18 years	VEG	350	70	150	75	75	30	250	35	—	—	30	—
		NON	350	50	150	75	75	30	150	40	30	30	30	—

1. As recommended by the nutrition expert group of the Indian Council of Medical Research, 1971.
2. In place of ground-nuts an additional 25 g or 30 g of fats and oil may be included in the diet.

Source: Gopalan and Narasinghrao. *Dietary Allowances for Indians*. Indian Council of Medical Research, Special Report Series No. 60. Nutrition Research Laboratory, Hyderabad-7, p. 87–90; compiled from Tables 23 to 28.

Religion prohibits Hindus from eating beef and Moslems from eating pork. Buddhists of India live strictly on a vegetarian diet whereas those of Burma, Thailand and some other Asian countries are non-vegetarian. The Hindu population consists of both vegetarian and non-vegetarian sections. There is a blending of ecological, religious and socio-economic factors in the evolution of the pattern of diet among different categories of Hindus. The tendency towards vegetarianism is deep-rooted, with the availability of milk as an important additional factor. Milk is a major item of the vegetarian diet and the only source of animal protein. Table 50 shows the distribution pattern of milk in India in relation to the vegetarian and non-vegetarian diet.

Table 50
Distribution of milk and vegetarianism and non-vegetarianism among Hindus

Status	*Per capita* milk (in ounces)	Diet pattern of Hindus
I. *Northern India*		
Punjab and Haryana	14.2	Mostly vegetarians
Rajasthan	8.7	
Uttar Pradesh	8.2	
II. *Southern India*		
Mysore	3.7	Brahmins, vegetarians; non-Brahmins, mostly non-vegetarian
Kerala	1.4	
Madras	2.7	
III. *Eastern India*		
Orissa	1.8	Mostly non-vegetarians
West Bengal	2.6	
Assam	1.9	

Though the northern India Hindus of the plain are mostly vegetarian, the Hindus of neighbouring Himalayan States (Kashmir, Himachal), where milk-yield is low, are non-vegetarian. The followers of the Krishna cult of Sri Chaitanya in the non-vegetarian states of Orissa, Bengal and Assam are vegetarian. In these states it is common to find different members of the same family attached to different Hindu religious faiths and differing in vegetarian and non-vegetarian eating habits. Hindu widows live on vegetarian food. They also exclude from their diet such foods as garlic and onion which historically have been thought to have sex-stimulating properties. Local variations occur in the inclusion or exclusion of these articles in the diet. Taboos and superstitious ideas as well as likes and dislikes about food are prevalent in all cultures. Some Indian tribes eat snakes and lizards, but other tribes dislike these foods. Frogs are plentiful in India and though they are a popular dish in Europe and China, they have never been considered a food item in India. The Chinese are said to be more versatile in their diet pattern.

Data on the level and quality of diet for the various people of the world have been made available by the Food and Agriculture Organization of the United Nations. Table 51 shows *per capita* food supplies and their nutrient content for Asian countries and some other selected countries of the world. The protein and fat components of the diet are very low in Asia. Among the countries listed, the total calorie value of the diet

Table 51. Per capita food supplies and their nutrient content in Asian and some selected countries

Country	Period	Cereals	Potatoes & other starchy foods	Sugar and sweets	Pulses nuts & seeds	Vegetables	Fruit	Meat	Egg	Fish	Milk	Fats and oils	Estimated nutrient content: Total proteins	Animal proteins	Fats	kc
		grammes per day														
Asian countries																
Afganistan	1961–62	476	1	9	2	64	69	37	2	—	223	2	68.4	15.8	29.7	2 050
Iran	1960	394	10	52	11	22	101	44	5	2	176	18	59.6	13.4	37.2	2 050
Iraq	1960–62	355	5	81	15	156	196	55	3	2	207	10	60.7	16.8	32.7	2 100
Israel	1964–65	278	98	106	27	307	385	128	60	18	391	49	85.8	39.7	95.5	2 820
Jordan	1964	320	31	52	31	444	855	21	9	2	122	54	59.0	9.6	69.7	2 390
Lebanon	1965	332	35	117	34	288	472	84	12	6	295	36	74.2	25.1	66.4	2 530
Syria	1963	429	22	40	25	180	404	40	4	2	231	36	71.7	16.4	64.3	2 360
Turkey	1960–61	611	105	28	36	288	340	37	5	6	193	22	97.5	15.9	53.9	3 110
Sri Lanka	1965	356	88	49	78	105	26	5	5	15	57	10	44.5	7.9	43.9	2 080
People's Republic of China[1]	1967	387	247	10	39	82	82	47	8	10	9	7	57.0	12.2	—	2 050
India	1964–65	404	37	50	61	—	45	4	1	3	123	11	53.9	5.7	26.8	2 110
Indonesia	1961–63	350	329	19	22	—	41	14	3	13	2	13	38.2	4.5	31.2	1 980
Japan	1965	394	173	50	43	293	98	28	24	76	100	19	77.6	24.6	45.1	2 350
Pakistan	1964–65	457	27	48	17	37	75	10	1	4	200	16	50.7	9.5	32.2	2 260
Philippines	1965	365	120	50	16	75	129	36	7	45	42	7	49.7	15.9	26.6	2 070
Thailand[1]	1963–65	412	91	29	52	108	128	28	9	19	13	4	46.9	9.9	—	2 140
Other selected countries																
Sweden	1965–66	191	260	119	8	97	235	140	32	57	703	67	81.5	53.7	137.1	3 000
United Kingdom	1965–66	213	282	137	17	162	156	203	42	26	290	62	88.8	52.7	143.1	3 250
Canada	1965–66	185	197	139	16	218	213	233	40	19	638	55	94.8	63.0	141.6	3 130
United States	1965	182	123	133	22	268	225	273	49	14	657	60	92.0	65.1	146.1	3 140
Argentina	1964	366	169	93	8	100	221	242	17	8	301	42	84.9	47.9	103.0	3 100
Peru	1963	235	358	75	26	229	189	69	6	20	108	18	57.3	18.3	42.4	2 090
Egypt	1963–64	586	40	46	29	281	242	36	4	14	124	20	84.1	12.6	49.0	2 930
Australia	1964–65	233	118	142	12	184	227	291	33	15	585	40	90.2	60.0	130.5	3 160
New Zealand	1965	236	174	123	13	234	195	304	47	19	747	55	108.6	73.6	154.8	3 460
U.S.S.R.[2]	1964–66	428	378	106	19	186	58	106	19	28	476	33	—	—	—	—

1. Data from Diet Atlas of India, National Institute of Nutrition, Indian Council of Medical Research, Hyderabad, 1971, Table XXVI.
2. *FAO Production Year Book*, 1971.

Source: The State of Food and Agriculture. Rome, FAO, 1967. Compiled from Annex Tables 8A, 8B and 8C, p. 167–77.

also falls below the FAO nutritional target for the world (1965–66) in Afghanistan, Iran, Sri Lanka, China, India, Indonesia, Pakistan, Philippines and Thailand. Fig. 87 shows the protein, fat and total calories of Asian and other selected countries. These figures show only national averages and do not show dietary differences among people of different economic strata of society.

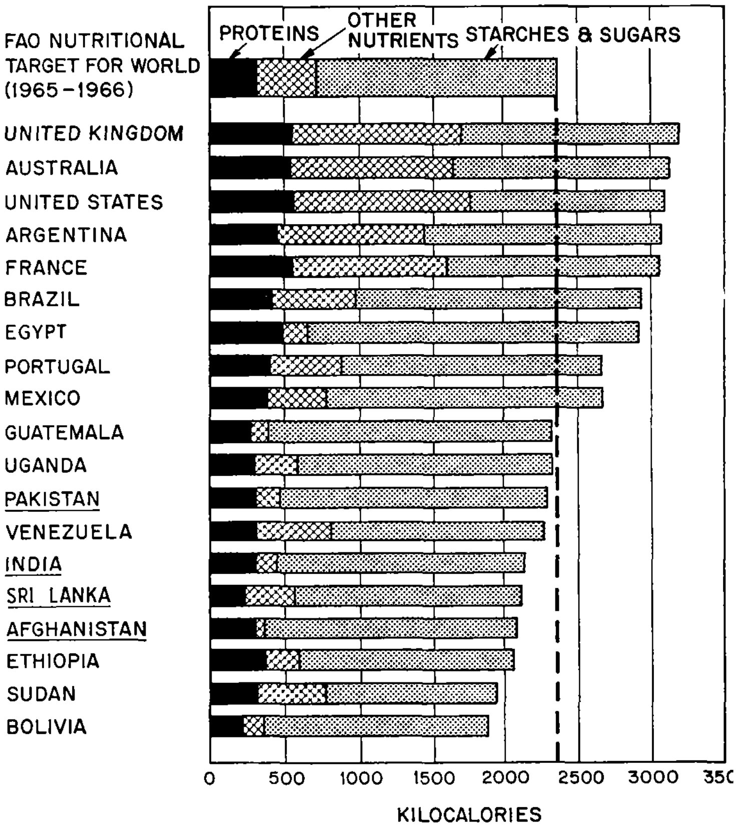


Fig. 87
Diet levels in Asian and other countries

Sources: The State of Food and Agriculture. Rome, FAO, 1967. Compiled from Annex Tables 8A, 8B and 8C, p. 167–77.
Diet Atlas of India. National Institute of Nutrition, Indian Council of Medical Research, Hyderabad, 1971, Table XXVI.
FAO Production Year Book, 1971.

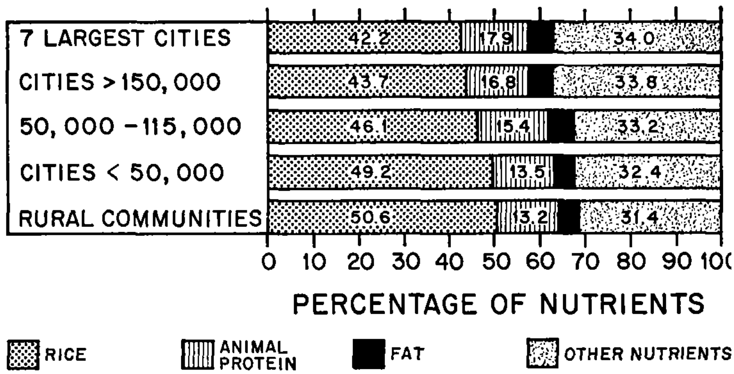


Fig. 88
Percentage of nutrients, 1969

Source: National Survey of Nutrition by Bureau of Public Health, Ministry of Health and Welfare, Japan.

The findings of the national survey of nutrition of Japan (Fig. 88) reveal that, in that country, the dietary pattern does not vary widely among various populations. The people living in the seven largest cities consume 8.4 per cent less rice, 4.7 per cent more animal protein, 1.1 per cent more fat and 2.6 per cent more 'other nutrients' than the rural population.

Undernutrition

Undernutrition is poor nourishment of the body as a result of an inadequate quantity of food. The energy value of the diet is the limiting factor in undernutrition.

In undernourished persons, there is loss of fat and protein from the body and, in severe cases, an accumulation of water. During the initial phase of undernourishment, water loss from the body is considerable. Gradually the amount of loss decreases and reaches a point after which there is water retention. The water content in severely undernourished infants is about 85 per cent whereas in normal infants it is around 65 per cent.

Undernutrition stunts growth. Infants born to underfed mothers are underweight. This is reflected in the birth weight of infants born to mothers of different socio-economic levels, as shown in Table 52.

Table 52
Mean birth weights according to socio-economic status

Place	Population	Subjects	Mean birth weights (grammes)
Madras	Indian	Well-to-do	2 985
		Mostly poor	2 736
South India	Indian	Wealthy	3 182
		Poor	2 810
Bombay	Indian	Upper class	3 247
		Upper middle class	2 945
		Lower middle class	2 796
		Lower class	2 578
Calcutta	Indian	Paying patients	2 851
		Poor class	2 656
Congo	Bantu	Very well nourished	3 026
		Well nourished	2 965
		Badly nourished	2 850
Ghana	African	Prosperous	3 188
(Accra)		General population	2 879
Indonesia	Javanese	Well-to-do	3 022
		Poor	2 816

Source: World Health Organization. *Nutrition in Pregnancy and Lactation*. Tech. Report Series, No. 302. Geneva, 1965, p. 14.
Lowenberg et al. Food and Man. New York, Wiley, 1968, p. 164.

It is obvious that this difference in birth weight is due to the difference in nutrition of the foetus and is not of genetic origin.

The average birth weight of infants in developing countries has likewise been found to be lower than that in advanced countries. During the first six months of life

the weight curves of both groups of children run almost parallel. But the weight curve of the majority of children of the world belonging to developing countries shows little increase in weight between the ages of 1 and 3. This is due to a condition of severe undernourishment. Children in the rural areas of many developing countries are breast-fed up to the age of about 2 years. Breast milk is sufficient to support growth for the infant's first six months, but beyond that the amount of breast milk becomes insufficient and supplementary food is needed. Unfortunately, there is little 'weaning' food such as cow's milk available to feed them. Even when available, the families are often too poor to purchase the food in sufficient amounts.

In urban areas, the breast-feeding period tends to be shorter. The children are often bottle-fed with diluted milk preparations, and the lack of knowledge by mothers of nutritional needs fosters undernourishment.

The underfed children are also more susceptible to diarrhoea than the well-fed ones ('weaning diarrhoea'). This is the primary cause of infant mortality in developing countries.

Undernourished persons lack energy and consequently the capacity to work; they are vulnerable to various diseases. Undernutrition not only affects physical changes but also causes changes in social behaviour and influences mental development. There is a distinct possibility of mental damage to pre-school children as a result of undernutrition. In this regard, the importance of the pre-school years is emphasized by the fact that the brain achieves nearly 80 per cent of its total growth by four years of age.

Undernutrition may occur due to non-availability of food, incapacity to purchase it or voluntary reduction in the diet.

Table 53
Regional *per capita* requirement and availability

Regions	*Per capita* estimated minimum calorie requirement[1]	*Per capita* calorie available[2]	Nutritional status
Group I		2 150[(a)]	Undernourished
Far East (including China)	2 300		
Near East	2 400		
Latin America	2 410		
Africa	2 340		
Group II		3 060[(b)]	Well nourished
Europe (including U.S.S.R.)	2 590		
North America	2 590		
Oceania	—		
World (FAO nutritional target 1965–66)	2 400	2 420	Just in balance

1. Excluding Argentina, Uruguay and Paraguay.
2. Including Argentina, Uruguay and Paraguay.

Sources: (a) *Oxford Economic Atlas of the World*, fourth edition. Oxford University Press, 1972, p. 20.
(b) P. V. Sukhatme. *Feeding India's Growing Millions*. Bombay, Asia Publishing, House, 1965, p. 80.

Within historical times innumerable famines have appeared in Asian and European countries. The last of the great famines, the Bengal famine of 1943, caused at least 1.5 million deaths. Famines still occur sporadically in many parts of the world.

The present extent of undernutrition in the world can be seen in Table 53 which shows the nutritional status and income level of the regions of the world.

Undernourished regions of the world contain about two-thirds of the world's population. The main factor in the regional disparity of nutritional conditions is the imbalance between the percentage of the world's population and income.

It must be underlined that the figures given above are regional or national averages. As available food is never distributed equally among the entire population, the extent of undernutrition in poorer sections of human populations of countries is likely to be wider and deeper than shown.

Malnutrition

Malnutrition is poor nourishment of the body resulting from an inappropriate (usually deficient) supply of essential nutrients.

Undernourished individuals get too few calories to maintain normal body weight and normal activity. Malnourished persons, on the other hand, may get enough calories, but wrong proportions in the quantities of nutrients such as building and regulative foods. Undernourished people are malnourished also, but the reverse is not necessarily true.

Malnutrition is by no means confined to undernourished regions of the world, but it is prevalent also in regions where there are seasonal variations of food supplies. Poverty and ignorance of the basic principles of nutrition are major causes of malnutrition.

A selected list of common deficiency diseases is shown in Table 54.

A pilot study made in India between 1956 and 1962 showed that the prevalence of goitre cases was reduced from 40 per cent to about 17 per cent by the use of iodized

Table 54
Some common deficiency diseases

Nutrient	Effect of deficiency
Carbohydrate	Loss of weight, acidosis
Certain fatty acids	Skin disorders (eczema)
Protein	Kwashiorkor, anaemia
Vitamin A	Xerophthalmia, blindness
Thiamine (B_1)	Beriberi
Nicotinic acid	Pellagra
Riboflavin (B_2)	Vascularization of cornea, circum corneal infection, chelosis, angular stomatitis
Vitamine C	Scurvy
Vitamin D	Rickets
Vitamin E	Sterility, eye abnormalities
Vitamin K	Failure of blood clotting
Calcium	Rickets
Iron	Anaemia
Iodine	Goitre

salt. One major problem about its use is that in tropical countries the crude and moist salt used makes the iodine compound unstable.

The prevalence and pattern of some nutrition deficiencies is due to dietary habits. Pellagra is essentially seen in maize-eating people of Latin America, parts of Africa, Egypt, Portugal, Yugoslavia and some areas of India. Maize is deficient in the vitamin niacin and this deficiency causes pellagra.

Beriberi is common in the rice-eating countries of the world. Symptoms of beriberi include edema (swellings), pain and loss of use of the limbs and heart failure. Thiamine, whose deficiency causes beriberi, is present in whole rice, but it is lost in milled or polished rice. Thus, the people using polished rice usually suffer from beriberi. In the par-boiled rice, used mainly in southern India, the vitamin is withdrawn into the centre of the grain and is not lost, even after milling. Home-pounded or unpolished rice also retains this vitamin. In 1958, beriberi was considered by the World Health Organization (WHO) to be a serious health problem in a number of rice-eating countries in southern and eastern Asia, especially Burma, Thailand and Vietnam. The disease was found to cause many deaths in breast-fed infants between the first and sixth months of life. Surveys in Burma and Thailand by WHO officers also supported these conclusions. There was some evidence that the prevalence of beriberi increased in association with the spread of small rice mills in a number of rural rice-producing areas in eastern Asia. It was common in Japan in the early decades of the century, but it became a rare disease there as a result of changes in the national diet and preventive measures taken by the authorities.

A prolonged deficiency of vitamin A leads to xerophthalmia (atrophy of the conjunctiva) and eventually to complete blindness. Xerophthalmia is now considered as one of the serious health hazards of the world. It is present at all ages, but young children are affected most. Xerophthalmia is most prevalent in South-East Asia, especially in Sri Lanka, southern India, Burma, Malaya and Indonesia. It also occurs in Latin America and the Middle East, but with a lower frequency.

Fig. 89 shows the distribution of some vitamin deficiencies and Table 55 shows the percentage of some deficiency diseases in Indian children.

Protein-calorie malnutrition

Protein-calorie malnutrition (or PCM) is a condition of malnourishment resulting from deficiency of protein which is invariably accompanied by a deficiency of calories. It is a condition of undernourishment as well as malnourishment. Protein deficiency affects mainly development and growth, so infants and children are affected most by PCM.

Kwashiorkor is a severe form of PCM. In India it was formerly called 'nutritional distrophy'. Kwashiorkor is prevalent in children between ages 1 and 4. A child with kwashiorkor has retarded growth, puffy legs, distended abdomen and a 'moon face' due to edema or retention of water in the tissues. Skin lesions are noticeable. There are changes in colour and texture of hair. The liver is enlarged and other internal organs are affected. Psychomotor changes are marked in the child's peevishness and mental apathy.

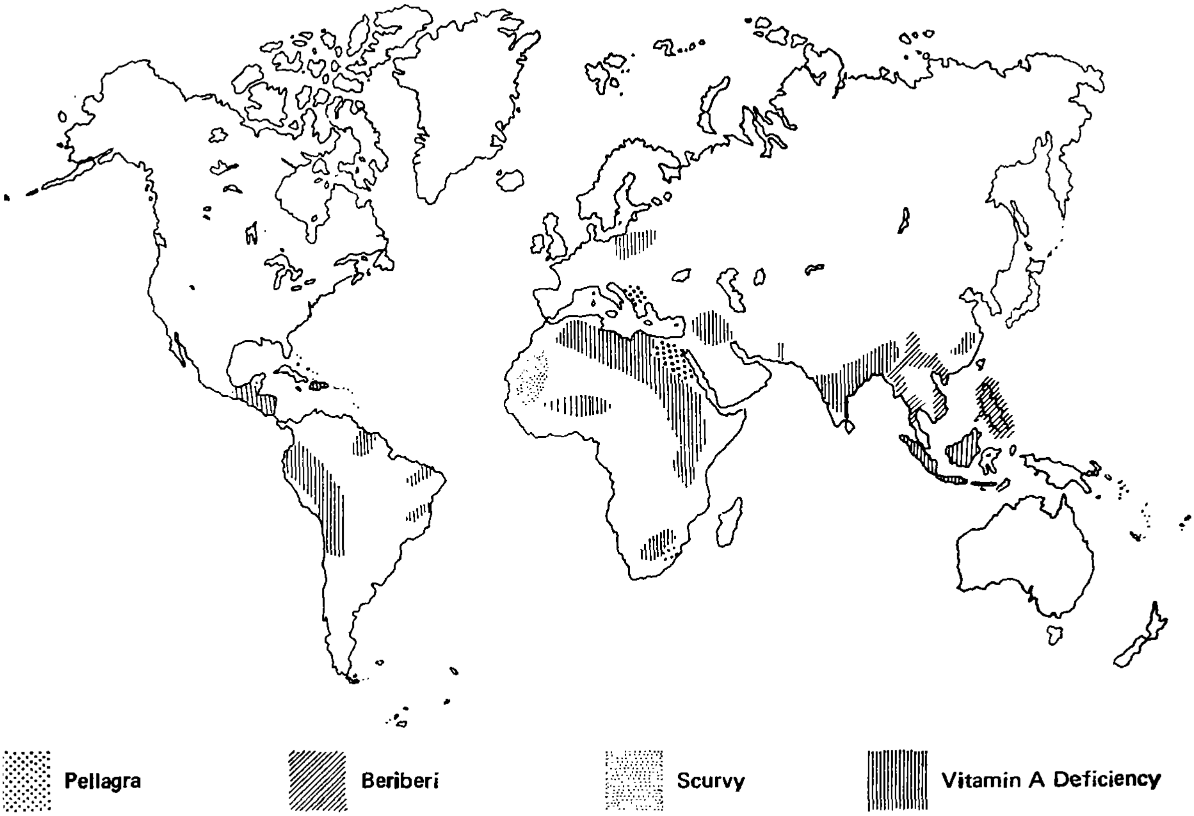


Fig. 89
Geographical distribution of some vitamin deficiencies
Source: World Health Organization. *Activities in Nutrition, 1948–1964*. Geneva, 1965, p. 27.

Table 55. Nutritional deficiency in Indian children[1]

Deficiency disease	Uttar Pradesh 1954–56	Madras 1954	Hyderabad 1954–56	Bombay, Saurashtra 1954	West Bengal 1954–56	Bihar 1954–56
			(in percentage)			
Bitot spot	5.0	4.4	2.2	0.8	—	9.8
Angular stomatitis	3.0	1.6	4.9	14.5	6.0	7.7
Glossitis	—	2.6	3.0	12.0	6.3	10.1
Phrynoduma	0.5	3.7	—	—	—	7.6
TOTAL	8.5	12.3	10.1	27.3	12.3	35.2
Total number of children examined	798	2 293	5 419	12 410	2 040	2 519

1. Three per cent of children were suffering from kwashiorkor.

Source: R. P. Mishra. *Medical Geography of India*. New Delhi, National Book Trust, 1970, p. 167.

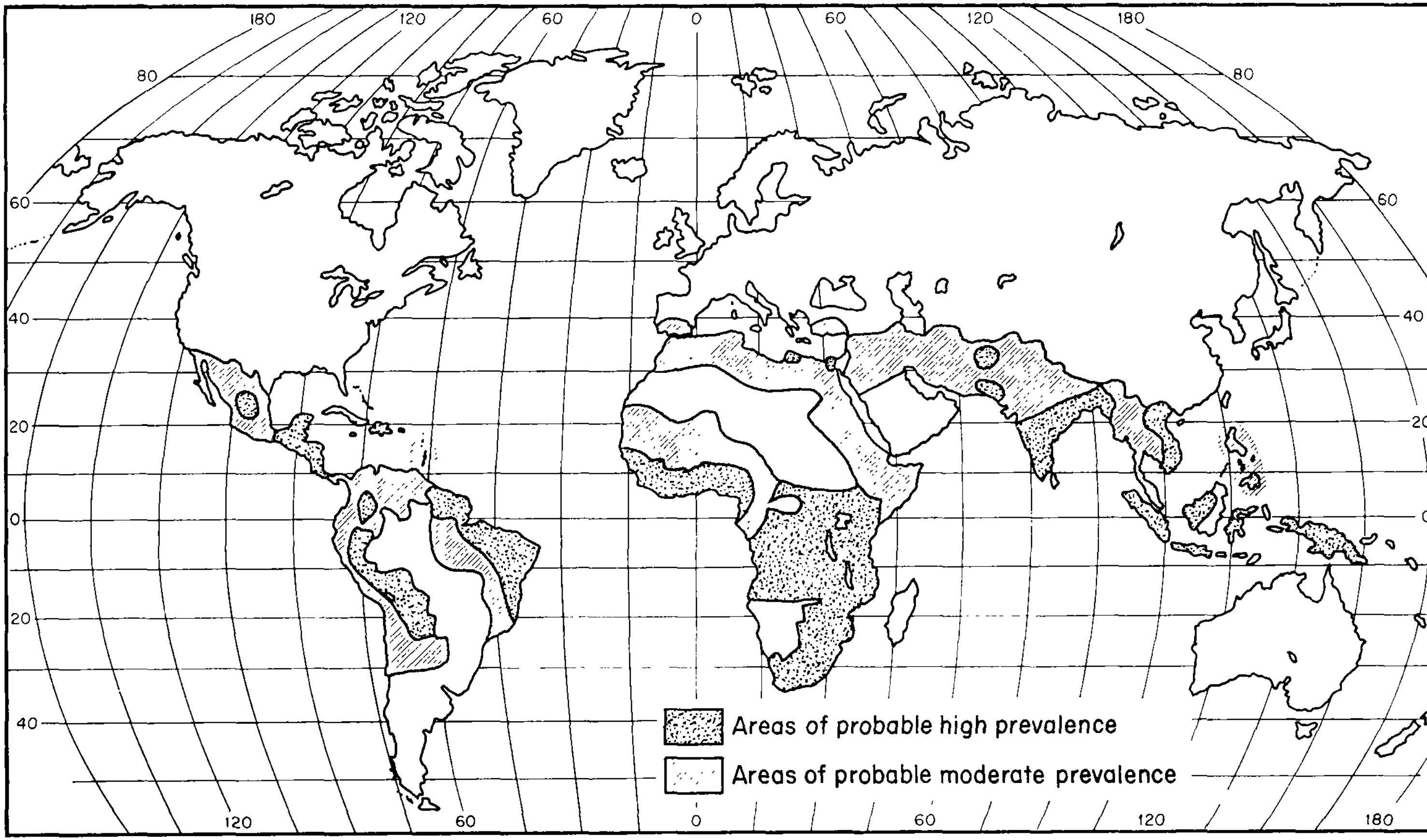


Fig. 90
Geographical distribution
of protein-calorie malnutrition
Source: Infant Nutrition in the Subtropics and Tropics. Geneva, WHO, 1968.

The newly weaned child is usually fed a starchy food as a supplement to or a substitute for breast milk in the countries of Asia, Africa and South America. Usually the starch is made into a gruel with water before it is fed to the child. The gruel is made from cassava or yam (low in protein) in the starchy-root-eating regions of Africa, from steamed plantains in Buganda, from sweet potatoes in southern Asian countries, from corn or maize in Latin America and Egypt and from rice in India and China.

Studies supported by WHO in different countries of the world including Ghana, Guatamala, India, Jamaica and Uganda, have shown that protein malnutrition is the result of interaction of several factors such as lack of protein, deficiency of calories, infectious diseases, ignorance of the principles of good nutrition and inadequate purchasing power.

Community surveys in southern India showed that 1–2 per cent of the pre-school children had kwashiorkor.

Fig. 90 shows the geographical distribution of protein-calorie malnutrition.

Storage, mishandling and poisoning of food

Storage of foodstuffs is essential to meet food requirements of both rural and urban populations for the entire year. Cereals, pulses and nuts are dry and thus easy to store. Fleshy fruits and vegetables, on the other hand, are very perishable and difficult to store. Providing a supply of such foods out of season requires preservation.

All human foodstuffs are perishable. The causative agents of their deterioration are micro-organisms such as bacteria, yeasts and moulds, and chemicals (ferments). Micro-organisms are the chief cause of food spoilage. Since water is necessary for the growth of micro-organisms, foods like fleshy fruits and vegetables which contain a high percentage of water are spoiled more readily than dry foods.

In one form or another preservation of food has been practised since prehistoric times. Vegetables and fruits were commonly preserved by drying; fish and meat by drying or salting; cereals by drying; fruit juices as wine or cider; and milk as cheese or ghee. Thus drying, salting and fermenting were major methods employed for preservation. With increased understanding of causes of food spoilage, techniques of food preservation have progressed more rapidly. The methods now commonly used for preservation of food are freezing, drying, canning, fermentation, pickling, as sugar concentrate and preservation with chemicals. Atomic radiation is also now used to preserve food.

Pickling and drying are very common in Indian homes and other tropical and subtropical countries. In pickling, food is preserved in common salt or vinegar or lime juice; spices and oil may also be added. Mustard oil, rape seed oil and sesame oil are generally used in India. Among the pickles, mango pickle is most common, followed in order by cauliflower, turnip, lime, chili and bamboo. (Use of bamboo pickles is now being discouraged, because the use of young bamboo shoots as pickles or other

forms such as *hendua* or dried powder from the shoot is depleting this plant in western Orissa.)

In tropical and subtropical regions preservation of fruits and vegetables by sun drying is practised during the dry season. Some important fruits preserved thus are apricots and grapes in Afghanistan, dates in Tunisia, Egypt, Algeria and Saudi Arabia; jack fruits and mangoes in India, peaches in Australia, southern Africa and California. Ripe bananas are also dried. Prunes, persimmons, cherries, berries, papaya, guava, tomato and brinjal are also dried to some extent. The dried tomato and brinjal serve as the main components of vegetable dishes in western Orissa during times of scarcity. In villages of South-East Asia including Eastern India, fish is preserved by drying through heating.

In canning, a syrup of sugar or corn is used as a covering liquid for fruits, whereas brine is used for vegetables. Chemical preservatives are sometimes used to stop or retard fermentation, acidification or other decomposition of food. The two important preservatives are benzoic acid (including benzoates) and sulphur dioxide (including sulphites). The containers used for canned food are made of glass or tin. Plastic containers are now becoming popular as well.

An additive is a substance which is added to food to serve some useful purpose. It may protect against spoilage or it may retain or enrich texture, appearance, taste or flavour.

Enrichment of food by the introduction of deficient nutrients is now considered as the best method of preventing certain nutritional deficiencies. Polished rice is impregnated with thiamine and other nutrients in definite proportions to get 'enriched rice'. Salt is iodized and recommended for people deficient in iodine.

Mishandling of food may be considered as improper handling of food materials resulting in lowered nutritional quality of food or in the poisoning or spoilage of it.

Thiamine, which is necessary for prevention of beriberi, is present in the germ and outer coats of rice, but it is lost in polished or white rice. The total loss due to milling in rice-eating Asian countries is of great importance. An additional problem is that rice is cooked in large amounts of water. Often the excess water is thrown away usually after cooking and the water soluble vitamin is thus lost. In rural areas the farmers keep this liquid over night and use it, with the starch granules it contains, as a drink for breakfast.

Washing vegetables already peeled for cooking also removes soluble vitamins and mineral salts. Valuable food is also lost by discarding the leaves of vegetables such as radishes or onions.

Food poisoning may be caused by chemical or bacterial agents or by the consumption of poisonous plants or animals. Chemicals such as metals from containers or injurious preservatives may poison the food. Grey enamelled utensils of poor quality may contain antimony. Fruit acids may dissolve the antimony and cause poisoning. Cadmium-plated utensils cause cadmium poisoning. Zinc-containing utensils also may cause poisoning.

Parasites or their cysts such as tapeworms and muscle worms and bacteria or their toxins also poison food substances. These food substances may be infective or

toxic. Foods like meat, fish, eggs and milk may be contaminated by the bacteria of *Salmonella* group and produce acute gastro-intestinal irritation, i.e. vomiting, diarrhoea and pain in the abdomen. These bacteria are non-sporing and are destroyed by heat but the endotoxins can withstand a temperature of 100° C. Some outbreaks of food poisoning traced to ice creams, milk, pastries, cakes, etc. are considered to be caused by *Staphylococus*. Botulism is a neurologic disease. It is caused by neurotoxins produced by *Clostridium botulinum* in underprocessed canned foods. The organism is common in the soil, but is anaerobic. Thus if oxygen is absent, as in canning, it can grow in a variety of foods. Botulism is now very rare due to modern methods of food-processing which eliminate the heat resistant *C. botulinum*. However, it is still occasionally caused by soil contamination of food or improper canning.

Many people are aware that some mushrooms, fish and shellfish are poisonous. Symptoms of poisoning have occurred in Japan after eating oysters and asari (a type of shellfish). The Minamata disease was due to high mercury concentrations in sea-water and in marine edible animals. (For poisoning by pesticides, see Fig. 64.)

Food resources

Man procures his food from the plants and animals living on land and in water (oceans, lakes and rivers). The surface of the earth is roughly 50,000 million hectares in area. Twenty-nine per cent of this area, about 14,000 million hectares, is land and the rest is covered by water. Out of about 3,000 million metric tons of food that is produced annually, about 2 per cent comes from water in the form of fish. The rest comes from land. Food from land comes mainly from agricultural land which is only 10 per cent of the total land area: 51 per cent of this cultivated land produces cereals; 1.6, potatoes; 1.2, sweet potatoes and yams; 4.5, pulses and 8.6, oil seeds. These plus sugar constitute 66 per cent of the total food by weight. Vegetables, fruits and nuts constitute another 16 per cent of the food and these also come mainly from arable land. The remaining 18 per cent is contributed by meat, milk and eggs. Table 57 shows the distribution of

Table 56. Land area, population and food production in Asia, China and India as percentage of world total

Regions	Land area	Cultivated land	Population	Cereals, pulses potatoes, sweet potatoes, yams, oil seeds, sugar[1]	Meat milk eggs
World	100	100	100	100	100
Asia	20.80	31.19	55.87	35.0	15.4
China	7.35	7.66	21.83	17.0	3.5
India	2.45	11.35	15.45	8.2	4.4

1. Rough estimation.

Sources: FAO Production Year Book, 1971.
Unesco Statistical Year Book, 1971.

Table 57. World food production[1] (in million metric tons)

Production of major crops, 1971	World	Europe	N. C. America	S. America	Asia	Africa	Oceania
Total cereals	1 309.0	203.2	293.6	54.0	284.6	65.5	15.0
Potatoes	306.4	134.7	17.6	8.5	13.6	2.7	1.0
Sweet potatoes+yams	147.7	0.1	1.5	3.2	16.0	23.5	0.2
Cotton (lint)	11.8	0.2	2.9	0.8	2.6	1.3	—
Total pulses	44.8	3.3	2.8	3.1	14.5	4.8	0.1
Total oil seeds	112.0	5.0	42.5	7.1	18.1	8.7	0.3
Sugar: centrifugal	72.3	15.8	15.8	9.8	11.0	5.2	3.2
non-centrifugal	11.3	—	0.5	1.1	9.1	—	—

Production of vegetables, fruits and nuts, 1970	World	Europe	N. America	L. America	Far East +Near East[2]	Africa	Oceania
Total vegetables	252.116	58.454	22.020	11.388		4.368	1.316
Total fruits	221.127	71.022	21.410	42.588	43.722	23.320	2.350
Total nuts	3.387	1.129	0.297	0.181	1.468	0.311	0.001

Major livestock products, 1971	World	Europe	N. America	L. America	Far East +Near East[2]	Africa	Oceania
Meat in dressed carcass wt.							
Beef, veal & buffalo	38.862	8.644	11.098	6.472	2.308	1.955	1.505
Pork	35.577	12.334	7.190	1.796	1.980	0.281	0.236
Mutton (lamb and goat)	7.016	1.039	0.266	0.478	1.562	0.823	1.428
Total	81.455	22.017	18.554	8.746	5.850	3.059	3.169
Milk	400.932	150.431	61.698	23.896	55.407	9.792	12.358
Hen (eggs)	21.566	6.039	4.581	1.448	3.133	0.509	0.246

1. World total includes data of U.S.S.R. and China. Data of China not included in Asia or Far East.
2. Near East includes the African countries of Egypt, Sudan and the Libyan Arab Republic.

Source: FAO Production Year Book, 1971.

food production in the world on an item and regional basis. Table 56 shows comparative land areas, populations and food production in various regions of the world. Fig. 91 shows these data in a graphical form.

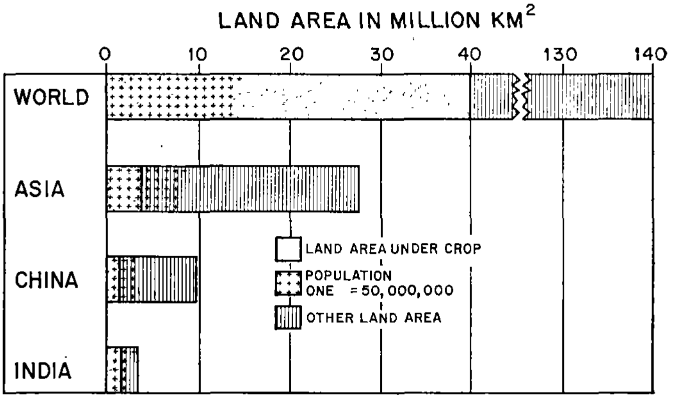


Fig. 91
Land area and population (1970) of world, Asia, China and India
Source: 1. *FAO Production Year Book,* 1971.
2. *Unesco Statistical Year Book,* 1971.

Asia possesses a little over 31 per cent of the world's cultivated land but has more than 55 per cent of the world's population. China and India, the two most highly populated countries of the world, together possess 19 per cent of the cultivated land but about 37 per cent of the population. Though India has more agricultural land than China, production of plant food is considerably less.

In modern agricultural societies, about 0.4 ha of fertile land is required to feed one person on a continuous basis. In Denmark, France and Sweden, food production self-sufficiency is achieved with about 0.4 ha *per capita*. In India and Pakistan, about 0.3 ha *per capita* is available for food production, and both are sometimes capable of self-sufficiency and at other times require food importation. The figures for China and the United Kingdom are 0.2 and 0.16 ha respectively, and both must import foods. The Philippines, Sri Lanka and Korea also have less than 0.2 ha *per capita*, and for Japan the figure is about 0.08 ha. For the United States, Australia and Canada the figures are 1.16, 1.60 and 2.6 ha respectively and each exports large quantities of food.

Clearly, as population density increases the land *per capita* decreases. Under these conditions, productivity must increase if *per capita* food supply is to remain at the same level. For example, before the Second-World War, the countries of Africa, Asia and Latin America exported 11 million tons of grain a year. In 1964, they were importing 25 million tons a year. Between 1959 and 1964, the population in Latin America increased 11.5 per cent compared with a food production increase of 6.5 per cent. Very probably this difference is a cause of the change from food export to food import conditions in Latin America.

From the above it follows that provision of adequate diets throughout the world will be very difficult unless population growth is slowed. In the meantime, it is imperative that every effort be made to increase food production and availability.

Some tropical and subtropical food sources

From the standpoint of dietary calorie production, sugar-cane is the most productive crop per ha per year being raised in the tropics, and world sugar-cane production has been on a sharp increase during the past twenty years or so. Other high calorie yield crops, in decreasing order, are oil palm, bananas, cassava, yams and sweet potatoes.

Annuals are generally not as productive as perennials since they utilize a smaller proportion of the year for growth. Partial compensation of this inadequacy can be achieved through the interplanting of crops, so that a second crop is growing even before the first crop is harvested. In this way, as many as six crops may be taken off the same ground in one year, though this is seldom done in practice. In India, about 13 per cent of the farm land is double-cropped. Of the annuals, cereals are generally highest in productivity with maize at the top. However, rice provides a higher sustained yield under continuous cultivation without fertilizers than most crops. This accounts for the tendency of high density populations to turn to rice for a continuous calorie source.

When dietary protein requirements are considered, crop priorities look quite different. Soybeans are very high in protein content but they have not become very significant in world diets, perhaps due to difficulty of preparation in a palatable form. Of the major cereals, rice is lowest in protein content. The protein of maize lacks two of the essential amino acids, tryptophane and lysine, and is low in a third, cystine.

Animal protein provides a more complete distribution of amino acids for human needs, but since this food is generated through secondary productivity, the efficiency

Table 58
Protein values of some foods

	Protein: percentage fresh weight	Protein: percentage dry weight	NPU
Plants foods			
Cassava	0.9	2.4	—
Potato	2.0	9.1	71
Maize (whole)	9.2	10.5	54
Rice (milled)	7.6	8.7	—
Wheat flour (white)	10.5	11.9	—
Apples	0.4	2.5	—
Spinach	2.3	29.1	—
Beans	22.1	25.0	—
Peas	23.8	26.7	—
Walnuts	15.0	15.5	—
Animal foods			
Beef	18.5	48.7	76
Cheese	25.0	39.7	69
Chickens	20.2	59.4	—
Eggs	12.8	49.2	95
Fish	18.8	72.6	75
Milk	3.5	27.6	90
Pork	15.5	46.8	79

Source: Don R. Arthur. *Man and His Environment*. New York, American Elsevier Publishing Co., 1969, p. 100.

of food production in quantitative terms is much lower than with vegetation. Fertilized fish ponds are the most highly productive of the animal sources in the tropics, followed in decreasing order by pigs, dairy animals, poultry and grazing animals. Some of these also compete with man for crops as food, so this fact must be weighed against the dietary protein contributed.

The utilization of protein from food sources can sometimes be estimated by feeding controlled diets to laboratory animals and determining bodily protein retention. This measure is termed the net protein utilization or NPU. Table 58 presents quantitative protein and NPU values of some foods. Note that some sources with a relatively low percentage of protein content provide a comparatively high NPU.

Present and future methods for increasing food availability

Most methods for increasing food availability depend upon producing more food, though some malnutrition problems could be alleviated if more effective food distribution could be worked out. For example, forest and savannah food exchange could be beneficial. Forest regions of the tropics can grow perennials, rice and interplanted crops, but very limited protein. Animal production is often difficult because of parasitism and limited grazing vegetation in forests. Savannah regions, especially with irrigation, are comparatively effective for non-rice cereal growth and cattle-raising. Interchange of food materials between these two types of regions should help meet the needs of both populations. Unfortunately, political, economic and social factors (such as eating habits) often mitigate against such adjustments.

Among the many procedures for increasing food production are the use of fertilizers, pesticides and irrigation. These were discussed elsewhere in this text and need not be repeated here. Intensive agriculture combining these techniques with the more efficient cultivation, harvestation and storage made possible through use of machinery and transportation system, may further increase productivity. Pitfalls in using these methods include erosion with mechanization, alkalization and salinization with irrigation, contamination of water supplies with fertilizers and accumulation of pesticides in soils, water and tissues of organisms.

Plant and animal breeding. Through selective breeding, varieties of crop plants are being developed with combinations of traits maximizing productivity. For example, strains of rice and wheat for tropical and subtropical regions have been developed which are stiff-strawed, short-statured, fast maturing, highly responsive to fertilizer and less sensitive to day length. A successful new strain of rice (called IR–8) matures in 120 days compared with 150 to 180 days for older varieties. The faster maturity and day-length insensitivity facilitate planting more corps per year and the dwarf stature of the crop plants results in greater efficiency in the conversion of raw materials into grain. With the introduction of these new varieties and the corresponding increased multiple cropping, production of these crops in some tropical and subtropical countries (e.g. Malaysia and Pakistan) has sharply increased. Massive use of such high yielding strains of rice and of fertilizers in Japan have contributed to an average rice yield over

three times as high in Japan as in India where largely unimproved strains are being used with very little fertilizer. Climatic differences between the two countries very probably also contribute to the difference.

Progress has also been made in animal breeding. Egg-laying by domesticated chickens has progressed from about 45 eggs per year to about 260 per year. Milk production by domesticated cattle has gone from 270 kg of milk per year to about 4,050 kg yearly in a modern agricultural society.

Supplements and synthetics. Several protein supplements have been developed. The procedure involves concentrating the protein component from sources known and accepted by the population to be served and insuring comparable, acceptable taste. 'Incaparina' is composed of yeast, calcium carbonate, vitamin A, corn and cotton-seed flour. It is about 25 per cent protein and can be diluted with water and drunk directly or added to other foods. The fact that it is a balanced protein product composed of plant materials is a major benefit since it is thus produced through the comparatively high efficiency of primary productivity. 'Laubina' is composed of vitamin A, vitamin D, bone ash, sucrose, chick peas and wheat. Fish protein concentrate is ground-up fish with oil and water removed. The powder is about 80 per cent protein. It can be produced from almost any kind of fish and it is tasteless, so it can be readily added to other foods as a protein supplement. The present product, obtained from fishing grounds off the coast of Peru, is used mainly as feed for livestock and poultry, but it could be used for human beings as well. Another variety of low-cost protein-rich food can be prepared from the remnant of oil seeds (soybean, cotton-seeds, peanuts, etc.) after the extraction of oils. This requires only the removal of the harmful component gossypol in cotton-seed and enzyme inhibitors in soya. These remnants are now used as animal feed or fertilizers.

Production of synthetic foods through modification of other organic materials is a future possibility. Soybeans can be treated with a mild alkali and processed to produce tasteless fibres of protein which can then be appropriately flavoured. The process is still rather expensive, but may hold some promise for the future. Micro-organisms which use petroleum for food could be harvested with a substantial yield of protein. One estimate suggests that only 3 per cent of the annual crude oil production of the world could supply adequate protein for the world human population using this method!

Increase of cultivated land. Table 59 summarizes the present cultivated area and potential arable area in the countries of Asia and the Far East. Considering the region as a whole, there is scope for a 50 per cent increase in the cultivated area.

But since population, land and water resources are not evenly distributed for efficient use and development in each country and because of limiting factors discussed earlier in the chapter, the total area under cultivation may be expected to increase from 298.1 to 316.6 million ha over a period of twenty years (Table 60), thus representing an over-all addition of 6 per cent.

On the other hand, population in the region may increase by 69 per cent during this time. It is, therefore, clear that enhanced food demand would have to be met not

Table 59. Present cultivated area and potential arable area

Country	Population (million)	Total land area (million ha)	Present cultivated area[1] (million ha)	Present cultivated area[1] (ha/*capita*)	Present cultivated area[1] (% of total area)	Potential arable area[2] (million ha)	Potential arable area[2] (% of total area)
Afghanistan	17.0	64.8	8.2	0.48	13	10.0	15
Brunei	0.1	0.6	0.1	0.58	17	0.2	33
Burma	27.7	67.8	16.8	0.61	25	40.0	59
Fiji	0.5	1.8	0.2	0.38	11	0.4	22
Hong Kong	4.2	0.1	0.0	0.00	25	0.0	25
India	554.6	326.8	170.0	0.31	52	194.0[3]	59
Indonesia	121.2	191.4	18.6	0.15	10	60.0	31
Iran	28.4	164.8	12.2	0.43	7	22.4	14
Khmer Republic	7.1	18.1	3.1	0.44	17	5.6	31
Korea, Rep. of	32.1	9.8	2.4	0.07	25	2.6	27
Laos	3.0	23.7	0.8	0.27	3	2.6	11
Malaysia	10.8	33.3	3.6	0.33	11	10.4	31
Mongolia	1.3	156.5	3.7	2.85	2	5.5	4
Nepal	11.3	14.1	2.4	0.21	17	7.2	51
Pakistan	136.9	94.6	29.4	0.22	31	31.4	33
Philippines	38.1	30.0	8.7	0.23	29	11.1	37
Singapore	2.1	0.1	0.0	0.00	25	0.0	25
Sri Lanka	12.6	6.6	2.0	0.16	30	3.5	53
Thailand	36.2	51.4	12.0	0.33	23	26.1	51
Viet-Nam, Rep. of	18.0	17.1	2.9	0.16	17	9.6	56
Western Samoa	0.1	0.3	0.1	0.53	33	0.1	33
Developing region	1 077.2	1 277.3	298.1	0.28	23	443.7	35

1. Including land under permanent crops, excluding permanent meadows and pastures; FAO figures increased slightly to allow for expansion up to 1970.
2. Assessed on the basis of reserves in permanent meadows and pastures, forested land and unused but potentially productive land.
3. Irrigation and power projects (five-year plans), Ministry of Irrigation and Power, 1967, p. 2.

Source: ECAFE. *Water Resources Series No. 40.* New York, United Nations, 1971, p. 32.

Table 60
Planned increase in cultivated and harvested areas (million ha)

	1970	1975	1980	1985	1990
Cultivated area	298.1	302.9	307.5	312.1	316.6
Harvested area	298.1	314.1	331.4	348.5	366.4
Cropping intensity (percentage)	100.0	103.5	107.8	111.7	115.9

Source: ECAFE. *Water Resources Series No. 40.* New York, United Nations, 1971, p. 33.

so much from an increased cultivated area alone but also from the increase in harvested areas and from more efficient utilization and management of land, soil, water and other resources responsible for agricultural production. Improved irrigation and drainage systems in addition to reclamation and flood control may further help to meet the increased food demand in the area.

Oceans. At present, most human food gathered from oceans is in the form of fish. The current annual catch from all of the world's oceans is around 45 to 50 million metric tons which is around 3 per cent by weight of the total annual food production of the world. During the past fifteen years, the world-wide fish catch has been increasing at a rate of about 5 to 7 per cent per year, doubling in less than fifteen years. Since about 1945, food production from the oceans has been increasing more rapidly than from terrestrial sources.

Probably the most outstanding example of this increased productivity is the development of the fishing industry in the region of the Humboldt Current off the coast of Peru. Since being opened up to fishing, this area has made Peru the most productive fishing nation on earth with about 20 per cent of the entire annual world fish catch.

It now appears that similar areas of high fertility occur in the Arabian Sea. If this resource can be properly developed, protein deficiencies of such countries as Sri Lanka, India, Pakistan, Iran, Saudi Arabia, Somalia, Kenya and Tanzania may be partly overcome.

The potential of the ocean is limited, however. Regions such as those mentioned above are comparatively localized, the open sea being much less productive. The nutrient-rich areas comprise only about 1 per cent of the surface of the ocean, but produce around 50 per cent of the annual world fish catch. Furthermore, it is thought that the North Atlantic Ocean is currently being fished about as extensively as possible and that future increases must come from other oceans, especially those of the southern hemisphere. It is estimated that the potential sustained yield of fish from the oceans is somewhere between 100 and 200 million metric tons. As present rates of fish-catch increase, this limit could be reached in little more than a decade.

Numerous developmental difficulties remain as well: in some areas, fuel costs for operating fishing vessels and cold storage plants would be difficult for local economics to bear; efficient equipment would also require extensive investment of capital. Necessary modification of eating habits of people to benefit fully from the available protein would require extensive education programmes in some cases. In addition, there is a problem of not knowing adequately what we are doing to the ecosystem structure and function of the ocean through an extensive fishing programme.

Additional possibilites of using the ocean as a food source are under consideration and experimentation. Harvesting and processing of seaweed as food would provide access to the primary productivity of the sea. So far, inefficiency in harvesting and general lack of acceptance of marine plant life as a major food source have limited the application of this approach. Various shellfish could be raised in estuaries, salt ponds, on rocky shores and in tank farms, but many natural areas are being damaged by pollution and additional marine farming (aquaculture) developmental research is needed.

11 Diseases and environment

Introduction

Man's interaction with other organisms is of many sorts. As a member of a biotic community, he is involved with them in nutritional relationships and in interdependencies which involve ecosystem factors such as the atmosphere and soils. He uses them for his own benefit through cultivation and domestication. He also has to fight various disease-causing organisms to preserve his health.

The causes of human diseases fall into several major categories and not all are the direct result of pathogens. There are many nutritional or deficiency diseases which result from inadequate quantity or quality of diet and have been discussed in the previous chapter. In addition, various illnesses result from accidental consumption or misuse of various toxins and drugs. There are also hereditary obstacles to normal health illustrated by such conditions as mongoloid idiocy and sickle-cell anaemia. (See Part I.) Mental and emotional stress represent additional areas of human diseases not caused directly by other organisms, though the human social climate may be a causative factor.

Superficially, disease and health may appear less directly involved with ecosystem structure and function than topics like climate control, soil maintenance, water purification and food production. However, among most organisms in nature, disease and death usually result not from inherent qualities of the organisms but rather from effects of factors which originate in the environment. Pathogens causing infectious diseases are of environmental origin as are conditions of undernourishment, vitamin deficiency, protein deficiency, numerous respiratory ailments, causes of accidental injury or death, and poisons. Diseased and healthy states of the body commonly represent certain relationships of the organism with the various biotic and abiotic conditions of the ecosystem.

Infectious diseases are those caused by the entry, subsequent development, and effects of other organisms in the (human) body. Various sources of infection of such diseases are summarized in Fig. 92. The means of transmission of these diseases varies, some being transmitted directly from one person to another, while others pass through one or more intermediate carriers called vectors which pass the pathogens on mech-

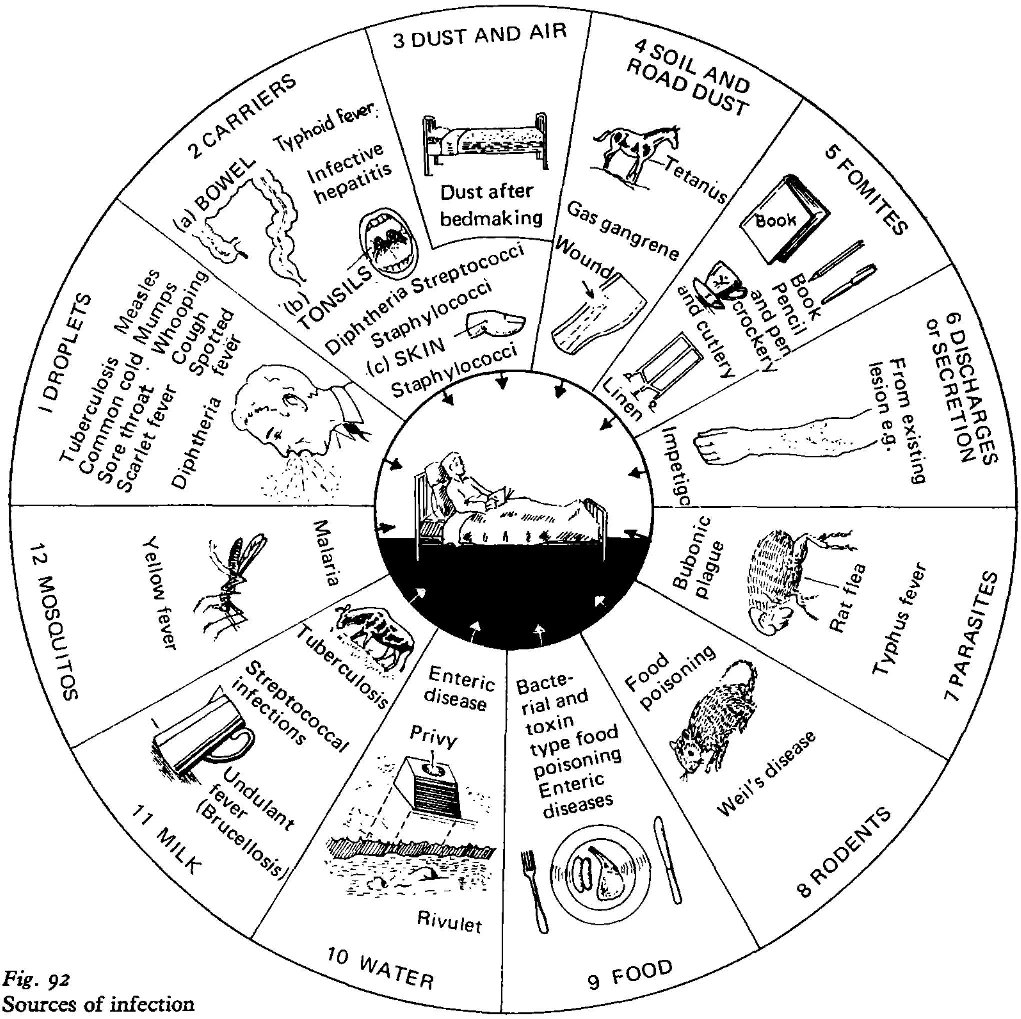


Fig. 92
Sources of infection

Source: N. Roper. *Man's Anatomy, Physiology, Health and Environment*. Edinburgh and London, Churchill Livingstone, 1973, p. 493.

anically by carrying them, or by their bite. Occasionally, environmental conditions may become optimal for the transmittal of pathogens. If a human population is susceptible to a disease, it may become widespread.

Pathogens and their transmission

Infectious diseases are caused by a wide variety of organisms. The major types involved are the viruses, rickettsiae, bacteria, spirochetes, protozoans, fungi and parasitic worms. These groups are discussed briefly below. Diseases known to be caused by viruses

include yellow fever, measles, smallpox, foot and mouth disease, rabies, mumps, poliomyelitis, influenza and encephalitis.

Rickettsiae are slightly larger than viruses, but also depend on living tissues for survival. Diseases caused include Rocky Mountain spotted fever, the tsutsugamushi disease (scrub typhus) and endemic and epidemic typhus.

Bacteria are the cause of infection of sores and of diseases such as cholera, diphtheria, pneumonia, tuberculosis and gonorrhoea.

Spirochetes are sometimes classed with the bacteria and sometimes with the protozoa. They are unicellular and very long relative to their width. Many of the smaller spirochetes are pathogenic. Some of the diseases caused by them are infectious jaundice, ratbite fever, syphilis, yaws, bejel and mal de pinto.

Protozoa are unicellular organisms which cause some of the most widespread of all tropical diseases. Organisms of this group cause malaria, leishmaniases, African sleeping sickness, Chagas disease and amoebic dysentery.

Fungi are most often involved as human pathogens in various skin diseases commonly referred to as 'ringworm'. However, they also are the cause of several serious systemic diseases such as histoplasmosis, coccidioidomycosis and *Torula* meningitis.

The parasitic worms include flukes, tape-worms and a number of nematodes. Flukes cause schistosomiasis and tape-worms are frequent parasites of the digestive system. A wide diversity of nematodes (hook-worms, pinworms, filaria, ascaris and trichina) occur as parasites in many parts of the body, e.g. trichina in various regions including the muscles, filaria in the circulatory system, ascaris and hook-worms in the bowels and liver flukes in the viscera.

Transmission of pathogens and the resultant spread of infectious disease are by one or more of the following mechanisms: by the air, entering the respiratory tract; direct contact; contaminated water and food; vectors. The viruses of measles, mumps and smallpox are examples of diseases transmitted by the air. In some cases, the transmission is only through direct human-to-human contact, as in syphilis (caused by a spirochete *Treponema pallidum*) and gonorrhoea (caused by a bacterium). These pathogens are normally passed from the infected individual to the uninfected through sexual intercourse. In schistosomiasis, cercariae (which form one stage in the life cycle of the fluke) escape from aquatic snails and, when encountering the human host by direct contact (while walking, swimming or even drinking water), they bore into the skin, enter the blood vessels, continue to develop, and eventually lay eggs in the small veins of the abdominal organs and intestinal walls. In yaws, a pathogenic spirochete is encountered by children through soil, especially through small wounds in contact with the soil. Various nematode infections, such as strongyloidosis and encylostomiasis, are also acquired when the pathogen bores through the skin. Contaminated water and food provide a special form of direct contact with pathogens. In amoebic dysentery (amoebiasis), the pathogenic protozoan *Entamoeba histolytica* in the human large intestine produces a resistant cyst; this cyst is carried out of the body with the faeces, it is readily transmitted to a new host by sewage-contaminated water, foods grown in soils fertilized with human faecal matter, insects or careless handling of foods. In the case of cholera, transmission of the infective agent is also by food and drink, the

Table 61. Some important tropical infectious diseases, causative agents, and vectors

Causative agent	Air-borne	Water-borne	Food-borne	Contact, skin penetration	Insect-borne	Other animal vectors
Virus	Common cold Influenza Bronchitis Measles Whooping cough Mumps Chicken pox Small pox	Poliomyelitis Infectious hepatitis			Yellow fever (mosquito) Dengue (mosquito) Encephalitis (mosquito) Sand-fly fever (sand fly)	Rabies (dogs) Encephalitis (ticks)
Rickettsiae	Q fever				Epidemic typhus (louse) Endemic typhus (flea)	Tick typhus (tick) Scrub typhus (mite)
Bacteria	Diphteria Pneumonia Meningitis Tuberculosis Leprosy	Typhoid Cholera	Typhoid Bacillary dysentery Amoebic dysentery	Gonorrhoea Tetanus Boils	Plague (flea)	
Spirochaetes				Syphilis Yaws	Relapsing fever (louse)	Relapsing fever (tick) Leptospirosis (rat)
Protozoa					Sleeping sickness (tsetse fly) Leishmaniasis (sand fly) Malaria (mosquito)	
Worms, flukes, parasites			Tape-worm Roundworm Thread worm	Scabies Jigger Hook-worm	Filariasis (mosquito) Loa loa (mango fly)	Tropical liver fluke (snail) Fluke (snail) Lung fluke (snail) Bilharziasis (snail) Guinea worm (water flea)
Fungi			Thrush	Ringworm Athlete's foot		

source of the pathogen is human waste and it is transmitted in water. Therefore, contaminated water, poor cooking habits, careless food handling and poor sanitation and waste disposal all facilitate transmission of the pathogen since the pathogen source is human waste and it is water-borne.

Vectors which transmit pathogens are very often (but not always) arthropods. In malaria, the mosquitoes of the genus *Anopheles* serve as the carriers of protozoans of the genus *Plasmodium*. The pathogens are transmitted from the salivary glands of the mosquitoes to the human host and infect first the liver and then the red blood cells. In yellow fever, the pathogen is a virus and the vector is a mosquito, *Aedes aegypti*. In tropical climates, this mosquito breeds in rain barrels, vases, and other containers usually associated with human populations. If the mosquito bites an infected human host, it becomes infective after about twelve days and remains a carrier of the virus for the rest of its life. It may thus infect any new host that it bites. In some forest areas, monkeys serve as mammalian hosts and the disease is maintained in the absence of man. In plague the pathogen is a bacterium that ordinarily infects rodents. It is transmitted from rat to rat, rat to man and man to man by fleas. In a severe human epidemic, the disease may be directly contagious from man to man. Table 61 summarizes the vectors and causative agents of the more important tropical infectious diseases.

Some ecological considerations

The spread of diseases is dependent upon various environmental factors. These factors are physical, social, cultural or combinations of these. Several examples are given below for illustrative purposes.

In schistosomiasis, physical factors as well as biological factors (besides the pathogen and host) are required. Certain species of aquatic snails are intermediate hosts, so the environment must have suitable aquatic habitats and it must include the snails. If the snails are missing, schistosomiasis is not transmitted. The same principle applies to vector-transmitted diseases in general, including the various rickettsial diseases transmitted by ticks, mites and fleas.

Measles outbreaks characteristically occur in the summer, both in the southern and the northern hemispheres. Thus, climate is again implicated, but the exact nature of its influence is not understood.

Cholera and amoebic dysentery are examples of diseases whose pathogens are transmitted through contaminated food and water. Environmental conditions facilitating such transmission often include high human population density, inadequate water and sewage treatment facilities, and inadequate sanitary regulations governing food handling and distribution.

Gonorrhoea and syphilis illustrate the influence of social and behavioural factors on disease transmission.

Cultural and technological conditions are also environmental factors which influence transmission of pathogens. Dam construction and deforestation indirectly

influence transmission of pathogens because of changes in the environment. Such environmental changes modify the entire ecosystem structure including relationships between pathogens and man.

Prevalence and distribution of some diseases

Wide variations in disease distribution patterns exist from country to country and locally. Thus, any attempt to describe briefly prevalence and distribution of major diseases on a world-wide basis will obviously result in considerable oversimplification. Nevertheless, it is possible to identify major regions of the world which are characterized by similar prevailing ailments and thus prepare a classification of disease regions. Fig. 93 presents four such regions and their distributions on a world map (on this map, the areas of countries have been adjusted to reflect their populations in 1950). Region 1 has temperate climates and about 800 million people. Region 2 contains

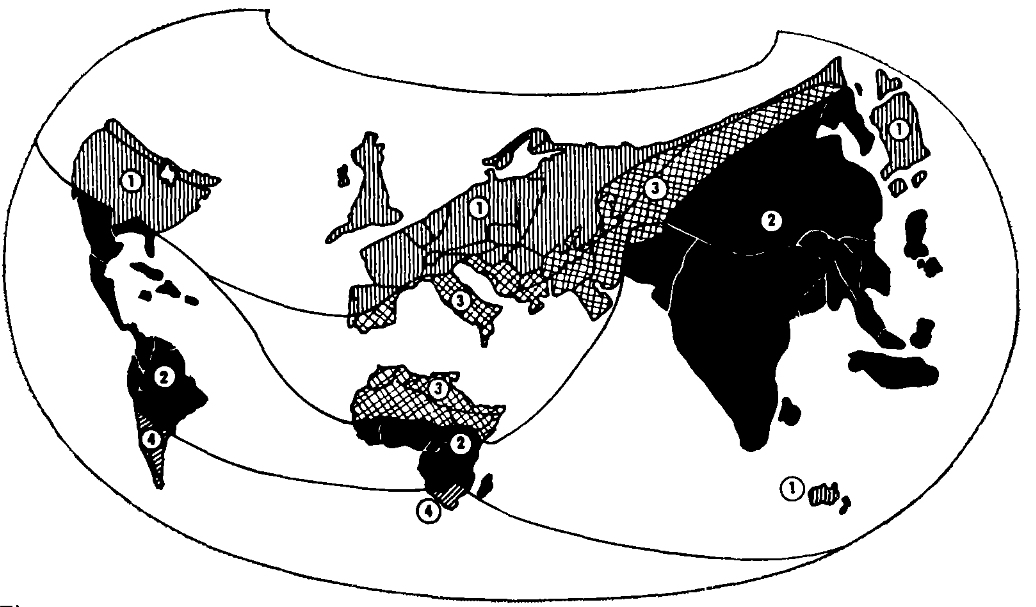


Fig. 93
World diseases by four regions
Source: Ralph Thomlinson. *Population Dynamics: Causes and Consequences of World Demographic Change*. New York, Random House, 1965, p. 114.

over 1,500 million people and Region 3 has about 400 million people. Region 4 is the smallest in land area and contains only about 100 million people. Table 62 presents prevalent diseases for these regions.

Table 62. Major disease regions of the world

Region	Prevalent ailments	Comments
1	Pneumonia, influenza, syphilis, tuberculosis, diabetes, mental disorders, heart diseases, cancer, common colds, suicide, industrial accidents, traffic accidents.	Healthiest part of the world. Preventive medicine important factor. Infectious diseases replaced by diseases affecting older persons. Many deaths caused by technology-related accidents. Largely temperate climate.
2	Malaria, amoebiasis, dysentery, yaws, syphilis, cholera, plague, yellow fever, sleeping sickness, smallpox, tuberculosis, leprosy, hook-worm.	Diseases largely parasitic and facilitated by climate. Poor sanitation often a contributing factor. Climate favourable for insect vectors formerly major factor, but now improved through use of pesticides. Largely tropical and subtropical.
3	Many of the above (from 1 and 2) plus contagious skin and eye diseases.	Intermediate between regions 1 and 2 and quite diverse. Trachoma very common.
4	Influenza, typhoid fever, tuberculosis, syphilis, flukes.	Substantial to excellent success available over influenza, typhoid, tuberculosis, syphilis.

Improvements in sanitation and the use of insecticides and preventive medicine have tremendously reduced infant mortality and infectious disease mortality. This is especially true in Region 1. In many regions, adequate sanitation is still lacking and modern drugs are unavailable.

Some causes and patterns of disease and death

Table 63 shows some of the commoner diseases with the number of deaths resulting from them in one country from America, Asia and Europe. In Sweden, the majority of deaths are due to pneumonia and heart diseases; both diseases of older people. Control of malaria in Sri Lanka appears more effective than in Peru.

Table 63
Mortality by diseases

Disease	Sri Lanka 1967 rate/100,000	Peru 1964 rate/100,000	Sweden 1967 rate/100,000
Tuberculosis	12.7	28.2	4.0
Typhoid	1.0	2.9	—
Cholera	—	0.1	—
Malaria	—	5.0	—
Smallpox	—	0.7	—
Typhus	—	1.6	—
Pneumonia	39.1	92.1	51.2
Enteritis, etc.	35.1	34.5	4.1
Anaemia	19.1	7.6	2.1
Heart diseases	51.9	19.2	345.7

Source: United Nations Handbook.

When societies succeed in controlling infectious diseases such as the above, the major diseases become degenerative diseases of old age and cancer. Moreover, mental diseases continue to be prevalent in many modern technological societies where infectious diseases have been largely controlled. Degenerative diseases of the heart and arteries, cancer, and various diseases of the nervous system are still largely unconquered, even with the aid of the best medical knowledge. The trend towards diseases of old age has further consequences in the process of demographic transition, which is discussed in detail in Part III.

In Fig. 93 most of Asia, with the exception of Japan, lies within Region 2 and is characterized by tropical parasitic diseases. Disease control is difficult as many complex factors are involved.

Data concerning causes of death are subject to many sources of error, especially in developing countries where reporting procedures may not be well developed. Caution in interpretation should also be exercised, since health problems vary extensively with local environmental conditions, thus deviating from data based on entire countries or continents.

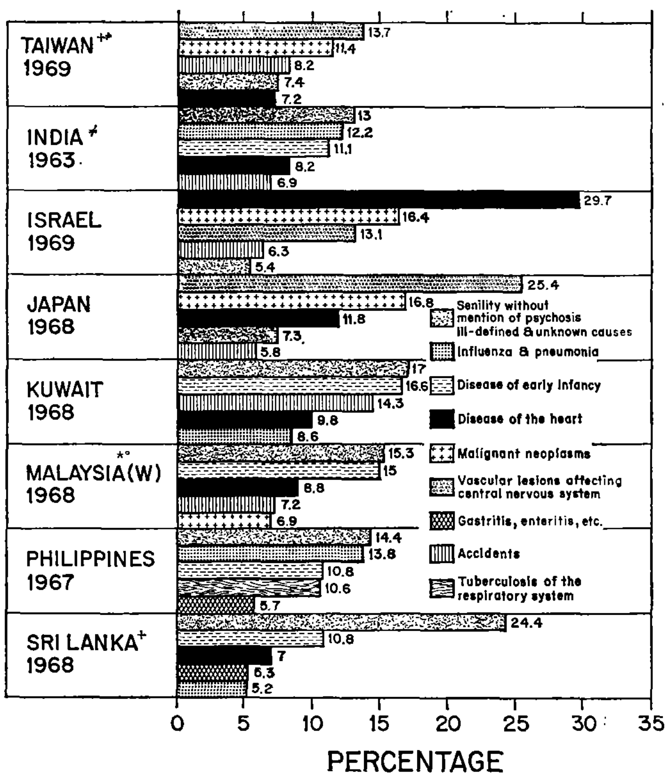


Fig. 94
Five principal causes of death, as a percentage of total deaths in some Asian countries and territories

Source: U.N. Demographic Year Book, 1970.

Note.

\+ Data tabulated by year of registration rather than by occurrence.
* Data exclude deaths of infants before registration of birth.
≠ Data on medically certified deaths in Poona and Bombay corporations and deaths in public hospitals of Rajasthan.
o Data for medically certified or inspected deaths only comprising 32.9 per cent of total deaths.

Fig. 94 gives five principal causes of death as percentages of total deaths in some Asian countries and territories. The prominence of 'vascular lesions affecting the central nervous system' in Taiwan and Japan contrasts sharply with the other areas listed. The same is true concerning 'malignant neoplasms' in Taiwan, Israel and Japan, and 'diseases of the heart' in Israel. The data for India came from two cities and from the public hospitals of Rajasthan which are located in urban areas. Thus, accidents as a major cause of death more closely reflect urban areas than the whole population.

Fig. 95 shows the percentage of deaths from infections and parasitic diseases in various Asian countries and territories. India and the Philippines show especially high figures while Israel and Japan have low percentages. In India, malaria has been endemic and considered the greatest health problem for years, though recent control efforts using pesticides have been highly successful. Smallpox, diphtheria, influenza, pneumonia, whooping cough, meningitis and leprosy are other prevalent diseases in India. Influenza and pneumonia are major causes of death. Intestinal infections and nutritional inadequacies are also prevalent. Table 64 gives the principal causes of death in the

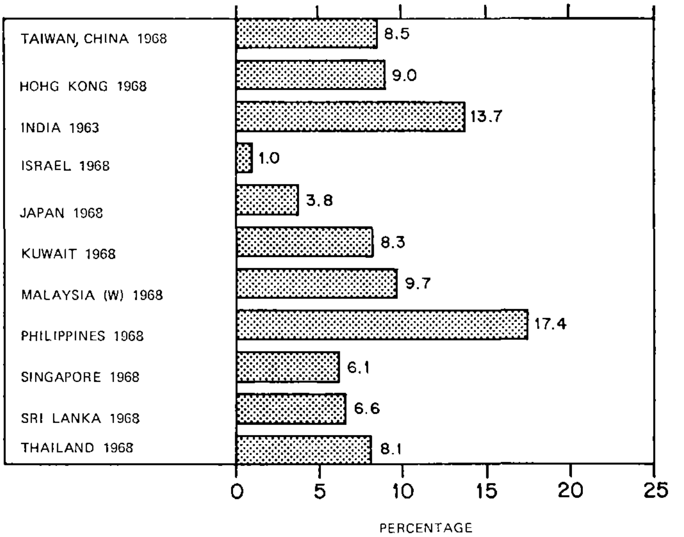


Fig. 95
Percentage of deaths from infectious and parasitic diseases in some Asian countries and territories

Source: 1. WHO. *World Health Statistic Report*, Vol. 25, No. 2, 1972.
2. *U.N. Demographic Year Book*, 1970.

Note.
India: Data on medically certified deaths in Poona and Bombay corporations and deaths in public hospitals in Rajasthan.
Kuwait: Percentage calculated from the data available from source 2.

Indian cities of Poona and Bombay and in public hospitals of Rajasthan State. In Taiwan, the main causes of death are changing quite rapidly. In 1946–47, over 2,200 deaths were attributed to cholera and 1,700 to smallpox. In 1952, diarrhoeal diseases were the most common cause of death, tuberculosis was third, malaria was among the first ten and six of the principal causes of death were communicable diseases. By 1962, only three of the first ten causes of death were communicable diseases. Even

so, enteric diseases in the mid–1960s were still in fifth position as a cause of death (even excluding diarrhoea in new-born infants).

Table 64
Principal causes of deaths in Indian cities of Poona and Bombay and public hospitals[1]

Cause	WHO ICD[2] Number	Total (%)
Senility, ill-defined and unknown causes	B45	13.0
Pneumonia	B31	12.2
Diseases of early infancy including infections, birth injuries, post-natal asphyxia and detelectosis	B42, B42, B44	11.1
Diseases of the heart	B25, B26, B27	8.2
Accidents	BE47, BE48	6.9
Gastritis, enteritis, etc.	B36	6.5
Tuberculosis of the respiratory system	B1	6.3
Malignant neoplasms	B18	4.0
Anaemia	B21	3.7
Vascular lesions affecting central nervous system	B22	2.2
Sub-total		74.1
All other causes		25.9

1. Public hospitals of Rajasthan State.
2. World Health Organization international classification of diseases, 1955, revision.

Source: U.N. Demographic Year Book, 1970.

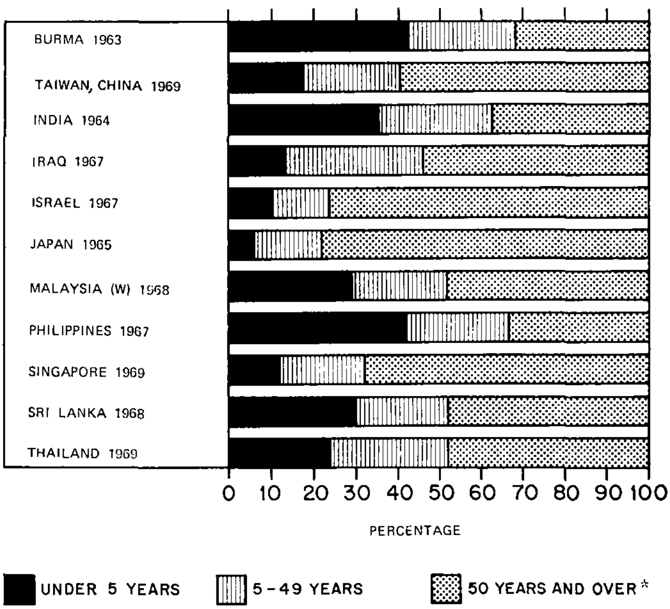


Fig. 96
Death by age group in some Asian countries and territories
Source: U.N. Demographic Year Book, 1970.

* Includes the death of unknown age group which is about 10 per cent for Thailand, 2.4 for India, 2.1 for Iraq, 1 for Singapore, 0.6 for Philippines, 0.3 for Burma, 0.1 for Malaysia, negligible for Israel, Japan, Sri Lanka and none for Taiwan.

Fig. 96 shows death percentage by age group in some Asian countries and territories.

Malnutrition does not normally appear as a major cause of mortality or morbidity, but its apparent absence is misleading. Medical authorities agree that malnutrition commonly contributes to death from other causes in developing countries. For example, the frequent deaths in early childhood from the 'weaning syndrome' which involves common diarrhoea, respiratory infections, and childhood diseases involving skin eruption like measles, are usually attributed to these latter symptoms. Nutritional deficiencies reduce body defences against infection, thus promoting greater incidence and severity of infection. One source suggests that 35 to 40 per cent of all deaths in Turkey may be due to malnutrition. Thus the role of malnutrition in developing countries is frequently significant though the extent of its impact may be obscured by other factors.

Data from developing countries show a pattern of high morbidity and mortality in infancy and early childhood. Fig. 96 illustrates this phenomenon compared to other age groups based on various countries of Asia. Fig. 97 provides a comparison of percentage of deaths of children under 5 years of age from several countries and territories

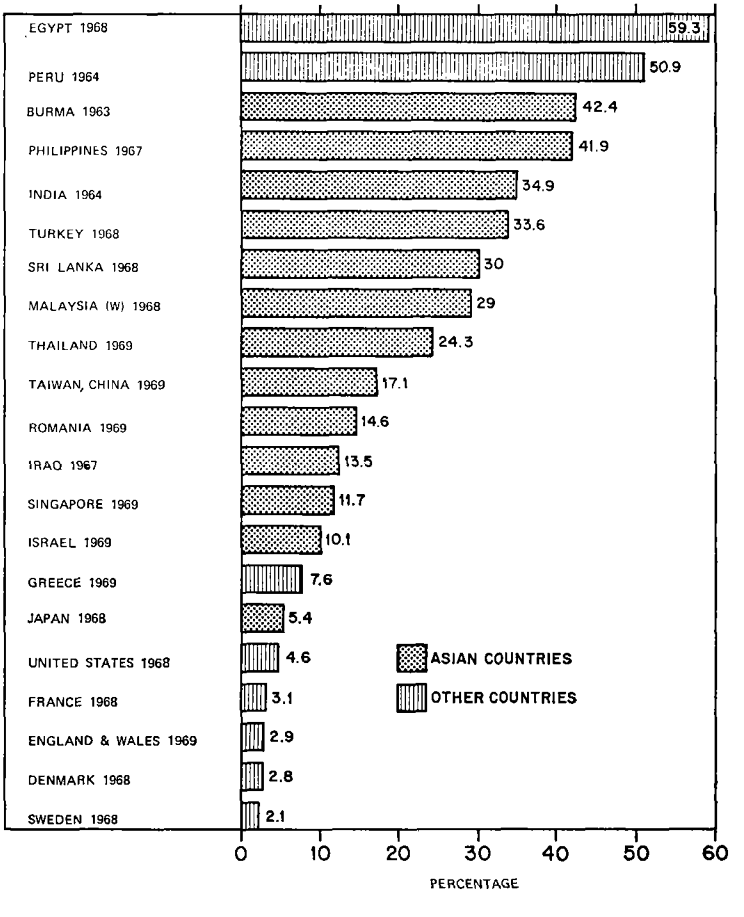


Fig. 97
Percentage of deaths of children under 5 years of age against the total number of deaths in selected countries and territories
Source: U.N. Demographic Year Book, 1970, p. 664.

of the world. Studies in Turkey have found that there are 125 to 253 infant deaths per 1,000 live births and that about 41 per cent of all deaths in cities and towns are made up of deaths at 0–4 years of age. In the Philippines, about 42 per cent of all deaths are children under 5 years of age.

Table 65 gives causes of infant deaths in India by percentage. Note that a total of 31.2 per cent of the deaths were attributed to a combination of intestinal intoxication, dysentery, and diarrhoea of the new-born. Table 66 shows causes of death in Japanese children; here the principal causes of death are congenital problems and accidents.

Table 65
Causes of infant deaths in India

Causes	Rate of death (in percentage)
Intestinal intoxication	22.6
Asphyxia	14.0
Prematurity	12.4
Bronchitis	7.5
Pyogenic infection	7.0
Nutritional maladjustment	5.4
Dysentery	4.8
Diarrhoea of new-born infants	3.8
Sub-total	77.5
All other causes	22.5

Source: Bhatia. A review of work on infant mortality in India. Part II. *Indian Council of Medical Research Special Report Series* (New Delhi), No. 32, 1957, p. 67.

Table 66. Causes of death of Japanese children under one year and from 1–4 years of age

First year		1–4	
Cause (by order of importance)	Rate per 100,000 births	Cause (by order of importance)	Rate (%)
Congenital malformation (S25)	202.4	Accident (BE47, BE48)	45.7
Birth injury, difficult labour and other anoxic and hypoxic conditions (S26)	194.2	Congenital anomalies (B42)	11.5
Pneumonia, bronchitis (S21, S22)	170.3	Pneumonia, bronchitis (B32, B33a, B46d)	11.5
Immaturity, unqualified (29)	131.7	Neoplasms (B19)	7.8
Other conditions of newborn (S32)	109.0	Enteritis and other diarrhoeal diseases, gastritis, duodenitis and chronic gastro-enteritis (B4, B46e)	3.9

Source: Vital statistics 1970. Japan, Health and Welfare Statistics Division, Ministry of Health and Welfare, 1971. Personal communication from Dr. Ohkura.

The compelling fact is that infant deaths are very high in many of the developing countries. Gastritis, enteritis and related digestive and intestinal illness which often accompany infectious and parasitic diseases, are common causes of death among the young.

Life expectancy

For every population there is an average life expectancy. It is governed, barring accidents and epidemics, by the quality of the environment and the general state of health of the individuals concerned. In Afghanistan, the life expectancy of anyone born between 1965–70 was about 37.5 years. In Iran, however, it is estimated to be 65 years for those living in rural areas and 62 years for those living in a city.

Members of societies, without the benefits of modern medicine, often have life expectancies at birth that hardly extend beyond 30 years. In such societies both the birth rate and infant mortality rates are high. About half the children die before they are 5 years old.

In developing countries, efforts are being made to reduce the infant mortality rate by skilled medical care and practice. Such care and practice have already made it possible to prolong the life of adults in many countries. Table 67 gives some figures for life expectancies of people in different parts of the world during the 1960s. It is to be noted that life expectancy is shorter in developing countries as compared to developed industrialized ones.

Table 67
Life expectancy of people in some parts of the world

Country	Life expectancy	
	Male (in years)	Female
Khmer Republic	44.2	43.3
India	45.23	46.57
Egypt	51.6	53.8
Japan	66.23	71.16
Canada	67.61	72.92
Sweden	71.56	75.35

Summary

Malnutrition and undernourishment are the underlying causes in decreasing resistance to infections which result in death. The broad groupings of 'respiratory infections' and 'digestive infections' are consistent major causes of mortality and morbidity, as are several of the viral infections. A general list of major disease conditions in many developing countries would include at least the following: tuberculosis, diphtheria, pertussis (whooping cough), pneumonia, bacterial dysentery, amoebic dysentery, typhoid, paratyphoid, yellow fever, measles, smallpox, poliomyelitis, influenza, typhus, cholera, leprosy, syphilis, yaws, malaria, Chagas disease, sleeping sickness or trypanosomiasis, cardiovascular diseases, cancer, undernourishment, malnutrition. Various diseases of early infancy continue also to be major causes of death.

Different diagnostic procedures may be significant in accounting for differences in

the prevalence of some diseases among societies. The prevalence of cancer and cardio-vascular diseases in technological societies is due at least in part to the fact that people survive to a greater age, providing more opportunity for malignant growths and organ deterioration to develop. Even here, ecosystem factors may be quite significant, however, since the existence of numerous cancer-producing substances is known, and an overweight condition due to excess food or inappropriate diet has strong implications in some cardio-vascular causes of death.

12 Mineral resources

Introduction

For the sake of convenience, natural resources can be studied under three categories:

1. Inexhaustable natural resources, e.g. the atmosphere and water in its cycle.
2. Replaceable and maintainable natural resources, e.g. water in a lake, soil, land (in its spatial sense), grasslands and wild animal life.
3. Irreplaceable or non-renewable resources, e.g. minerals and nature study areas.

This chapter concerns mineral resources which in their most general sense are taken to include non-living naturally occurring chemical compounds which are organic or inorganic in origin. They have been concentrated in the earth's crust by various chemical processes that are not necessarily controlled in any way by man's use of the end products. They include, besides metals, such elements as coal, petroleum, sulphur and phosphates as well as rocks themselves which are combinations of minerals. The earth contains only fixed amounts of these resources. They are products of millions or billions of years of earth's history. Unlike the renewable resources, they cannot be regenerated by rapid cyclic processes once they have been extracted from the earth.

Classification of minerals

Minerals have been classified in a variety of ways. For example, coal, iron and copper are sometimes referred to as basic and all others as contributory in terms of industrial economy; or certain minerals are called strategic because of their significance in national defence while others are non-strategic. Again, in terms of trade, they may be primary as products of first capture from nature, or secondary as products of recovered scrap or by-products. However, for the sake of convenience, all minerals can be divided into two main groups, metallic and non-metallic, each of which is further classified on the basis of their respective uses and physico-chemical properties as shown in Table 68.

Table 68. Generalized classification of minerals according to physical and chemical characteristics and use

					Example
Minerals	Metals	Ferrous	Iron		Iron ore
			Ferroalloying		Manganese ore
					Metallurgical chromite
					Molybdenum
		Non-ferrous	Base		Copper
					Tin
					Mercury
			Light		Aluminium
					Magnesium
					Titanium
			Precious		Gold
					Silver
					Platinum
			Rare		Uranium
					Radium
					Beryllium
	Non-metals	Mineral fuels	Fluid	Liquid	Petroleum
				Gaseous	Natural gas
			Solid	Hard coal	Anthracite
					Semi-anthracite
				Soft coal	Bituminous
					Lignite
		Other non-metals	Building materials		Sand and gravel
					Stone
					Cement materials
			Chemical materials		Sulphur
					Salt
					Chemical chromite
			Fertilizer materials		Phosphate rock
					Potash
					Nitrates
			Ceramic materials		Silica
					Clay
					Feldspar
			Refractories		Silica
					Fire clay
					Refractory chromite
			Abrasives		Sandstone
					Corundum
					Industrial diamonds
			Insulating materials		Magnesia
					Asbestos
					Mica
			Pigments and fillers		Ocher
					Clay
					Diatomite
					Barite
			Precious and semi-precious stones		Gem diamond
					Amethyst
					Amber

Source: William Van Royen; Oliver Bowles; Elmer W. Pehrson. *The Mineral Resources of the World.* Englewood Cliffs, N.J., Prentice-Hall Inc., 1952.

Earth's crust and ores

Most of our mineral resources are obtained from the earth's crust which is composed of about eighty-eight different elements. Various estimates of the composition of the earth's crust (shown in Table 69) indicate that there are only about eight or nine most abundant elements accounting for about 99 per cent of the total crustal mass. The remainder seventy-nine or eighty elements, therefore, occur only in traces but, nevertheless, they contain some of the most valuable mineral resources.

Table 69
Composition of major chemical elements in the earth's crust

Element	Earth's crust After Mason and Brian[1] 1952	Continental crust K. K. Turekian[2] 1969
Oxygen	46.60	45.2
Silicon	27.72	27.2
Iron	5.00	5.8
Magnesium	2.09	2.8
Aluminium	8.13	8.0
Calcium	3.63	5.1
Nickel	—	—
Sodium	2.83	2.3
Potassium	2.59	1.7
Titanium	0.44	0.9
Sulphur	0.14	—
Hydrogen	0.14	—
All others	0.83	1.0
TOTAL	100.00	100.00

Sources: 1. B. J. Skinner. *Earth Resources*. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1969, p. 22.
2. Guy-Harold Smith (ed.). *Conservation of Natural Resources*. New York, Wiley, 1958, p. 289.

These materials vary greatly in their relative geographical abundance and the degree to which they are concentrated in minerals. The rich local concentrations which can be exploited profitably are called ores. A mineral deposit may contain important quantities of an essential metal, but unless it is economically feasible to work, the deposit may be regarded as rock or at best as a marginal or submarginal ore.

Processing of an ore depends on various factors including: the rarity of the metal; the use to which it is put; the difficulty of extracting the element from the ore; the type of technology available.

As such, gold ores are processed with as low as 0.00016 per cent gold; copper one with as little as 0.9 per cent metallic copper and iron ore with 30 or more per cent iron.

The oceans and the atmosphere are big potential sources of certain elements. Their relative abundance is shown in Table 70.

Table 70
Relative abundance of elements in the oceans and the atmosphere

Element	Percentage composition of	
	Atmosphere	Oceans
Nitrogen	78.03	<0.002
Oxygen	20.09	85.79
Argon	0.94	(traces)
Carbon	0.03 (CO_2)	0.002 (C)
Hydrogen	0.01	10.67
Chlorine	(traces)	2.07
Sodium	(traces)	1.14
Magnesium	(traces)	0.14
Calcium	(traces)	0.05

Sources: R. F. Dasman. *Environmental Conservation*. New York, Wiley, 1972, p. 405.
Guy-Harold Smith. *Natural Resource Conservation*. New York, Wiley, 1965, p. 327–49.

Distribution and production of minerals

Different minerals are localized in different geochemical provinces. Their production is, therefore, limited to what exists in the earth's crust or in the sea and the atmosphere. It is further claimed that of all the quantities present, much has not been and may never be located by the technology known today. More than 1,600 are so far known. Among them about 50 are rock forming and about 200 are said to be of economic significance. Most, however, exist in minute quantities and are relatively unimportant. It is evident that some parts of the earth are relatively rich in mineral deposits and others poor. No single region is self-sufficient in all critical resources. North America, for instance, is rich in molybdenum but poor in tin and manganese; Asia is rich in tin, tungsten and manganese; Africa is poor in tungsten; most of the world's chromium and gold is in Africa (especially southern Africa); cobalt is also in Africa (especially Zaire) and the supply of mercury is essentially limited to European countries. However, most continents do have some share of the most abundantly occurring minerals in the earth's crust.

The present knowledge in most countries pertains mostly to easily accessible mineral resources. As such, these deposits have either already been mined or are in the process of being extracted in most of the industrially advanced countries, especially the United Kingdom and the United States of America. According to various reports, the United States has already changed from being a raw materials surplus nation to being a raw materials deficit nation. Among many other minerals, its demand for most of manganese, chrome, cobalt, tin and bauxite is met through imports. Foreign sources also help supplement its requirements of lead, zinc, tungsten, iron ore and copper. The present and near future sources of manganese are mostly in South America, Australia and Africa; the sources of tin mainly in South-East Asia; and most of the sources of aluminium ore in various underdeveloped tropical regions.

A recent Economic Commission for Asia and the Far East (ECAFE) survey of

mining and mineral resources indicates substantial gains during the twenty-year period (1945–65) in spite of unfavourable internal socio-economic factors affecting mineral resource development. The encouragement came mainly through increasing the share of responsibility on the part of national governments besides an over-all increase in both the world demand and the price of minerals. The review also points towards new discoveries of oil, gas and other minerals in various parts of the region, e.g. iron in Afghanistan, Korea and Malaysia; bauxite in Malaysia; bauxite, copper and manganese in Australia; tin in Laos; and mercury and copper in the Philippines. Such new finds are not uncommon in the world. Thus, the lack of production of certain minerals in certain areas is no proof for their absence. The problem perhaps lies in the lack of capital and proper know-how to undertake exploration and development of these resources. Possibilities, therefore, exist for developing countries to discover new sources and improve those already there by reorientation of governmental policies affecting mining industry in all its aspects, intensifying geological surveys and increasing the use of geophysical, geochemical and photogeological methods. As such, some countries in the ECAFE region have developed joint ventures with foreign capital and are consequently developing bauxite, lead, zinc, copper, iron and petroleum resources in Australia; copper, iron and nickel in the Philippines; tin, tungsten, antimony and iron in Thailand; and oil in Malaysia.

United Nations production indexes (using 1963 performance as the base, i.e. 100) indicate gains in all phases of mineral industry in 1969 compared with their 1968 performance. A similar upward trend is reported in the world production of most major minerals in 1970, the United Nations index of world mineral industry production for the extractive industries climbing eight points to 143. In the metals category, production increases were moderate to substantial. Of thirty-seven metallic mineral commodities, thirty registered gains in 1969 compared with 1968 results, declines were recorded for the remaining seven, the most significant being those for white arsenic (down 18.3 per cent) and primary smelter tin (down 2.7 per cent).

Table 71. World and Asian production of some non-ferrous metals, 1970–71 (in thousands metric tons)

Non-ferrous metals	World production		Increase/ decrease	Asian production		Increase/ decrease
	1970	1971		1970	1971	
Aluminium	10 213.6	10 884.3	+670.7	937.7	1 124.8	+187.1
Refined lead	3 987.6	3 866.5	−121.1	229.7	234.7	+ 5.0
Refined copper	7 544.2	7 317.0	−227.2	738.6	747.9	+ 9.3
Smelter zinc	5 077.3	4 967.1	−110.2	702.0	751.6	+ 49.6
Smelter tin	238.9	240.4	+ 1.3	119.0	119.3	+ 0.3
Smelter nickel	593.1	610.5	+ 17.4	89.9	103.0	+ 13.1
Magnesium	215.1	227.0	+ 11.9	10.3	9.7	− 0.6
Cadmium	16.38	15.23	− 1.15	2.57	2.71	+ 0.14
Mercury	10.03	10.74	+ 0.71	0.63	0.69	+ 0.06
Silver	9.40	9.00	− 0.40	0.47	0.50	+ 0.03

Source: World Metal Statistics, 1972.

World production statistics for some important non-ferrous metals during 1970–71 show increases for refined aluminium, smelter tin, smelter nickel, magnesium and mercury and decreases for others during 1971 over 1970 (Table 71). In Asia, however, only magnesium registered some decrease for the similar period. The rate at which the importance of iron has accelerated is very interesting. A century ago, iron ore production was only 4 per cent of what it is today. Its growth output was particularly rapid until 1910. The activity slowed down between 1910–40 to expand again since the latter date (Fig. 98). (See also Fig. 99.) Most additional iron resources since 1954 have been through increased developments in Australia, India, parts of Africa, Afghanistan, Korea and Malaysia where high grade ore containing about 60 per cent iron is extracted.

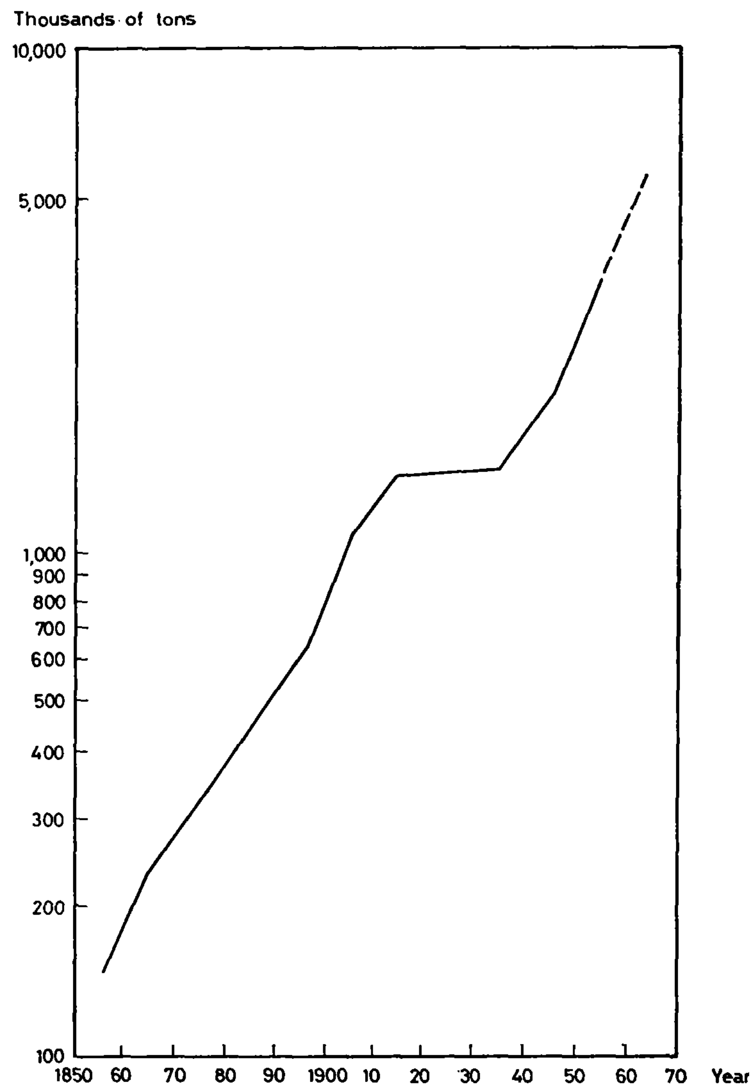


Fig. 98
Production of iron ore, 1855–1965
Source: United Nations Survey of World Iron Ore Resources: Occurrence, Appraisal and Use. 1970.

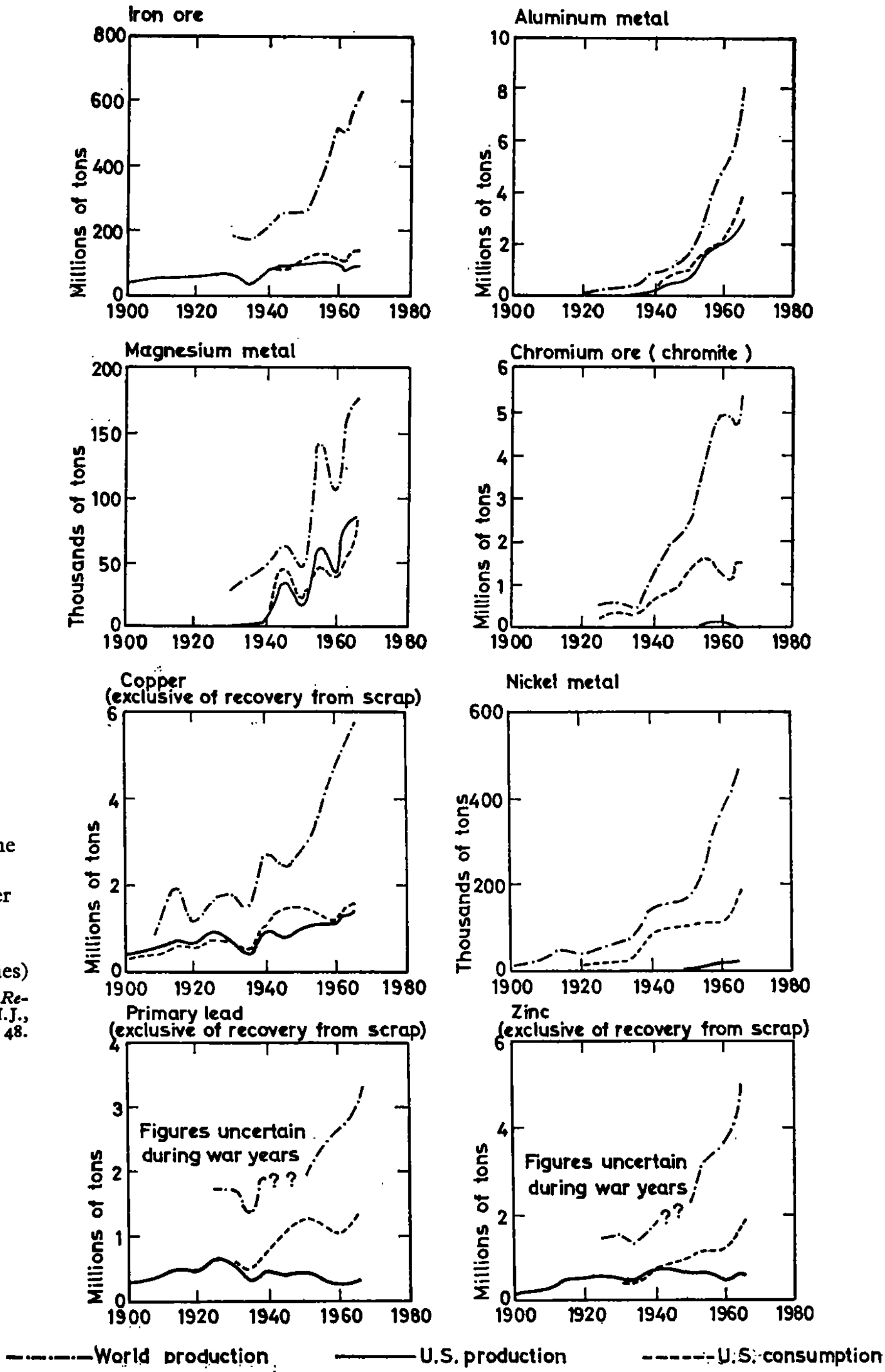


Fig. 99
World and United States production trends of some geochemically abundant and scarce metals together with the United States consumption rates (after U.S. Bureau of Mines)

Source: B. J. Skinner. *Earth Resources*. Englewood Cliffs, N.J., Prentice-Hall, 1969, p. 25 and 48.

Among twenty-three non-metallic minerals sixteen showed higher production levels in 1969 over 1968. Of those indicating declines, a 14.6 per cent decrease in graphite production was the most significant. On the other hand, most significant gains were reported for nitrogen content of nitrogenous fertilizers (9.7 per cent), gem diamond (8 per cent) and fluor spar (6.7 per cent). Industrial output as measured by the United Nations over-all production indexes rose 7 points to 157 for 1971 following a 10-point gain in 1969.

The mining industry in Asia and the Far East also reflects an over-all increase in spite of unsettled conditions in some parts. According to the twenty-year (1945–65) review, the pattern of production in certain countries showed minor fluctuations but declines in output were invariably offset by subsequent increases resulting in peak post-war production in 1965 (Table 72). The increase in mine production of minerals was followed by a comparable increase in value of total output of minerals or refined products. Most of these minerals were, however, exported to Japan to supplement its industries.

Table 72. Production of selected minerals in the ECAFE region in 1945 and 1965

Minerals	Production		Unit
	1945	1965	
Coal (excluding China)	75	191	million tons
Crude oil output	22	159	million kilolitres
Natural gas	45	23 433	million cubic metres
Iron	11	86	million tons
Manganese	371 000	3,000 000	tons
Chromite	95 000	889 000 (1957)[1]	tons
Tungsten	3 800	23 800 (1961)[1]	tons
Tin	11 000	128 800	tons
Copper concentrates	66 700	395 900 (1963)[1]	tons
Lead ore	173 800	518 100 (1963)[1]	tons
Zinc	183 700	825 500	tons
Mercury (excluding China)	126	1 263 (1960)	tons
Bauxite	120 000 tons	3.9	million tons
Gold	35 300	69 700 (1962)	kilogrammes
Silver	325 200	937 600 (1962)	kilogrammes

1. Peaks registered earlier as indicated in parenthesis followed by declines in 1965.

Source: ECAFE. *Mining Development in Asia and the Far East. A Twenty-year (1945–1965) Review*. New York, United Nations, 1967.

As stated in the United Nations Industrial Development Organization (UNIDO) monograph on industrial development (1969), the developing world's share in primary non-ferrous metal production ranged from less than 5 per cent for aluminium, 11 for zinc, 21 for lead, 40 for copper to 70 for tin.

Mineral consumption

As standards of living rise, mineral consumption does likewise. From the beginning of the industrial revolution, each generation is said to have consumed mineral resources equivalent to all previously consumed. In the United States, the consumption of minerals increased in the first half of the century four times as fast as the growth of its population. As such, it is the world's biggest consumer of materials: according to the Paley report, this country was using about half the world's supply and increasing its demand at about 3 per cent a year; the report further states that 'the country took out of the ground two and a half times more bituminous coal in 1950 than in 1900; three times more copper, four times more zinc, thirty times more crude oil'. The quantity of most metals and mineral fuels used in the United States since the First World War exceeds the total used throughout the world up to 1914. Although most materials are in heavily increasing demand, the hard core of the materials problem is 'minerals'. As a result, the United States and other developed nations are now increasingly turning to re-cycled metals, substitution and to foreign sources. By 1960 average annual *per capita* consumption of non-power minerals was said to be valued at U.S.$40 as compared with an average of less than $1 in the underdeveloped world.

World consumption of most mineral commodities has again advanced in the present decade both in terms of gross tonnage and on a *per capita* basis. The latter gains are less universally consistent but the former may be related to direct continued increase in population. One 1970 estimate points a general increase of 5 per cent per year in world mineral consumption. Another forecast expects a trebling of world demand by the end of the century.

A recent United Nations survey on iron states that during the short span of the past 115 years mankind has used about 16,000 million tons of iron ore, slightly more than half of this amount being consumed in the past twenty-five years. Of the high iron-consuming countries, Japan has recently achieved a prominent place among the world's steel producers, importing more than 90 per cent of all the iron ore for the purpose. During the fifteen-year period (1950–65) the apparent steel consumption of all developing countries rose from 9.7 million to 31.7 million tons crude-steel equivalent (Table 73).

Table 73
Apparent consumption, production and imports of steel in developing countries, 1950, 1960, 1965 (crude steel equivalent)

	1950	1960	1965
	(thousand tons)		
Apparent consumption	9 715	20 770	31 708
Local production	2 836	9 282	16 977
Imports	6 879	11 488	14 729
Local production as percentage of consumption	29.5	44.5	54

Source: United Nations Publication, Sales No. 68.II.E.4.

The total world and regional consumption and production of non-ferrous metals since 1931 are given in Table 74. In 1971 a total increase in world consumption is

Table 74
Bauxite and aluminium

Annual averages	Africa: Bauxite production	Africa: Aluminium production	Africa: Aluminium consumption	America: Bauxite production	America: Aluminium production	America: Aluminium consumption	Asia: Bauxite production	Asia: Aluminium production	Asia: Aluminium consumption
									In thousar
	—	—	—	70.6	7.4	7.1	—	—	
	—	—	—	363.6	50.9	38.7	—	—	0.9
	—	—	—	583.7	91.8	89.3	8.8	—	6.8
1931–40	0.2	—	0.1	801.0	136.3	112.8	139.2	12.1	24.1
1941–50	107.8	—	0.6	4 917.3	844.0	651.2	395.8	57.6	58.8
1951–60	571.9	12.6	6.9	10 616.7	1 889.5	1 491.6	841.7	101.9	116.2
1961–70	2 375.7	88.3	36.4	19 952.2	3 489.7	3 177.2	2 753.7	559.5	699.8
1969	3 188.1	159.8	60.2	25 154.2	4 565.2	4 095.5	3 416.6	860.0	1 192.2
1970	3 284.7	165.4	69.1	26 793.4	4 747.3	3 884.5	4 155.4	1 072.7	1 352.0
1971	3 553.6	191.2	75.0	27 931.7	4 774.9	4 333.1	4 332.9	1 269.8	1 448.0

Lead

Annual averages	Africa: Mine production	Africa: Smelter production[2]	Africa: Metal consumption	America: Mine production	America: Smelter production[2]	America: Metal consumption	Asia: Mine production	Asia: Smelter production[2]	Asia: Metal consumption
									Lead content i
		—	2.0		417.8	332.0		11.7	14.1
	39.4	10.7	5.0	567.2	553.5	489.1	27.4	24.6	34.2
	54.8	26.0	3.2	853.1	850.8	626.8	73.3	67.5	69.7
1931–40	33.9	23.7	5.7	772.0	713.1	435.5	101.7	89.6	111.6
1941–50	53.2	21.0	11.1	859.7	855.2	798.5	30.7	33.3	46.8
1951–60	209.9	68.2	20.7	865.0	972.1	830.2	120.4	103.5	119.1
1961–70	204.0	103.6	31.1	979.8	1 028.2	959.3	264.0	299.7	352.3
1969	208.6	137.1	34.2	1 171.4	1 203.8	1 093.0	291.7	370.0	467.3
1970	203.1	142.9	38.9	1 305.6	1 229.1	1 055.8	300.3	394.7	475.5
1971	206.2	131.9	43.9	1 319.0	1 157.6	1 112.0	323.8	409.7	480.7

Copper

Annual averages	Africa: Mine production	Africa: Smelter production	Africa: Metal consumption	America: Mine production	America: Smelter production	America: Metal consumption	Asia: Mine production	Asia: Smelter production	Asia: Metal consumption
									Copper content i
	8.3	—	0.7	513.4	487.8	257.3	40.2	36.7	18.0
	26.5	45.6	1.4	856.1	844.9	515.9	83.7	83.0	55.7
	107.6	95.5	3.3	1 056.2	1 053.5	644.1	72.3	66.8	88.6
1931–40	281.3	270.6	4.5	1 038.9	1 036.0	523.3	104.8	90.0	155.6
1941–50	422.7	394.0	18.6	1 584.2	1 610.9	1 314.4	83.9	76.4	89.6
1951–60	729.9	689.4	26.6	1 761.5	1 757.4	1 267.0	190.4	152.5	180.1
1961–70	1 119.6	1 093.5	42.6	2 569.7	2 418.3	1 725.9	367.6	440.3	666.9
1969	1 277.8	1 260.9	45.5	2 899.2	2 702.0	1 916.0	438.8	538.8	940.2
1970	1 286.7	1 281.7	55.3	3 157.8	2 830.0	1 816.8	478.9	658.2	988.4
1971	1 291.1	1 279.3	54.0	3 038.4	2 681.9	1 892.0	524.7	666.4	1 031.7

1. Including the U.S.S.R. in Asia.
2. As from 1951, production of refined lead.

Australia			Europe[1]			Total		
Bauxite pro-duction	Aluminium		Bauxite pro-duction	Aluminium		Bauxite pro-duction	Aluminium	
	pro-duction	consump-tion		pro-duction	consump-tion		pro-duction	consump-tion
aetric tons								
	—		136.2	9.8	9.9	206.8	17.2	17.0
	—		196.3	43.3	54.6	559.9	94.2	94.2
0.3	—	0.4	741.3	96.1	79.4	1 334.1	187.9	175.9
6.5	—	0.5	1 589.3	237.7	252.0	2 536.2	386.1	389.5
37.4	—	4.0	2 136.6	375.8	543.5	7 594.9	1 277.4	1 258.1
14.1	5.6	19.4	6 271.2	1 100.6	1 304.6	18 315.6	3 110.2	2 938.7
3 059.1	85.4	83.5	12 672.6	2 867.4	3 118.1	40 813.3	7 090.3	7 115.0
7 921.1	126.4	117.7	14 865.8	3 753.7	4 167.7	54 545.8	9 465.1	9 633.3
9 256.3	205.6	138.7	15 872.6	4 022.6	4 410.4	59 362.4	10 213.6	9 854.7
12 538.6	245.6	153.6	17 118.7	4 402.8	4 462.3	65 475.5	10 884.3	10 472.0

Australia			Europe[1]			Total		
Mine pro-duction	Smelter pro-duction[2]	Metal consump-tion	Mine pro-duction	Smelter pro-duction[2]	Metal consump-tion	Mine pro-duction	Smelter pro-duction[2]	Metal consump-tion
thousand metric tons								
	96.7	5.8		456.7	626.7		982.9	980.6
174.0	111.0	11.1	288.6	418.7	586.3	1 096.6	1 118.5	1 125.7
165.4	140.4	16.0	254.5	339.8	674.7	1 401.1	1 424.5	1 390.4
239.0	211.5	23.3	330.5	426.2	892.2	1 477.1	1 464.1	1 468.3
217.8	193.5	41.6	282.8	386.0	622.7	1 444.2	1 489.0	1 520.7
294.0	195.0	51.9	650.9	1 056.0	1 234.3	2 140.2	2 394.8	2 256.1
386.6	213.5	64.8	1 002.8	1 652.0	1 855.2	2 837.2	3 297.0	3 262.7
452.0	214.9	75.0	1 138.8	1 970.8	2 120.4	3 262.5	3 896.6	3 789.9
456.7	206.0	65.9	1 161.8	2 014.9	2 170.4	3 427.5	3 987.6	3 806.5
398.7	186.4	60.7	1 158.6	1 980.9	2 192.0	3 406.3	3 866.5	3 889.3

Australia			Europe[1]			Total		
Mine pro-duction	Smelter pro-duction	Metal consump-tion	Mine pro-duction	Smelter pro-duction	Metal consump-tion	Mine pro-duction	Smelter pro-duction	Metal consump-tion
thousand metric tons								
35.4	27.4	5.0	100.5	144.1	404.9	697.8	696.0	685.9
37.9	36.4	10.9	106.1	147.3	541.6	1 110.3	1 127.2	1 125.5
12.4	13.3	7.9	113.7	122.2	601.6	1 362.2	1 351.3	1 345.5
17.2	15.6	11.4	211.5	222.7	900.4	1 653.7	1 634.9	1 595.2
19.3	18.9	39.8	266.1	283.8	847.4	2 376.2	2 384.0	2 309.8
56.5	45.6	42.6	521.8	591.6	1 597.3	3 260.1	3 236.5	3 113.6
112.0	88.4	69.5	998.6	1 111.7	2 605.7	5 167.5	5 152.2	5 110.6
131.1	116.2	73.5	1 207.8	1 330.4	2 886.1	5 954.7	5 948.3	5 861.3
157.8	111.7	82.5	1 278.3	1 399.1	3 015.1	6 359.5	6 280.7	5 958.1
169.9	143.2	79.3	1 404.6	1 529.0	3 026.9	6 428.7	6 299.8	6 083.9

Table 74 (continued)
Zinc

Annual averages	Africa			America			Asia		
	Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption
									Zinc content i
		—	—		185.1	183.4		—	9.0
	23.2	—	0.8	514.0	422.4	337.9	34.9	18.7	21.1
	29.0	4.0	1.5	714.9	521.6	443.2	52.5	17.9	51.3
1931–40	24.5	13.9	3.7	815.1	561.3	439.4	73.0	44.6	94.8
1941–50	68.1	17.8	10.7	1 162.0	1 027.8	844.1	57.4	41.9	76.6
1951–60	235.2	60.5	20.3	1 273.9	1 162.9	954.9	213.8	155.9	187.9
1961–70	258.8	109.0	45.5	1 830.3	1 405.1	1 305.8	472.5	614.4	664.9
1969	250.4	126.0	56.3	2 234.1	1 608.7	1 502.0	554.6	921.4	917.9
1970	257.4	144.4	66.9	2 329.0	1 476.7	1 337.9	592.9	882.0	985.1
1971	253.1	162.1	67.8	2 365.8	1 322.3	1 403.6	626.0	941.6	995.4

Tin

Annual averages	Africa			America			Asia		
	Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consump- tion
									Tin content i
	0.9	—	0.4	16.2	1.2	40.6	74.8	78.2	5.7
	6.6	—	0.5	25.8	4.5	60.2	85.1	88.5	7.3
	9.2	—	1.3	33.8	2.8	74.1	101.8	103.0	9.9
1931–40	15.5	1.9	1.3	27.2	0.8	67.1	109.8	100.4	13.6
1941–50	27.4	8.3	1.9	39.7	31.2	69.2	67.8	53.5	10.8
1951–60	23.8	3.9	2.8	29.0	21.3	63.6	118.1	88.1	18.3
1961–70	20.5	11.3	3.1	29.4	8.4	68.5	129.7	123.0	42.2
1969	20.2	12.3	3.2	35.3	4.5	71.5	134.9	139.1	49.1
1970	19.8	11.6	3.1	37.5	9.7	65.9	140.3	142.0	48.7
1971	19.2	10.7	3.1	38.8	15.2	66.1	142.7	142.3	52.0

1. Including the U.S.S.R. in Asia.

Source: World Metal Statistics, 1971.

estimated to be more than 2.5 per cent over 1970 as against less than a 1 per cent increase in production for the same period. Consumption of tin in Asia has grown much less rapidly than that of most other metals where consumption has often exceeded regional production.

In 1971 for the second consecutive year, the world production of aluminium has exceeded consumption by over 400,000 metric tons, thus raising producers' stocks within two years by over 900,000 tons. This can be attributed to new smelters in the Federal Republic of Germany, the United Kingdom, the Netherlands, the United States, Japan and southern Africa. Production of smelter copper has also exceeded its consumption in 1971 though with a reduced margin.

Zinc smelter production in 1971 fell by 2.2 per cent but the consumption rose by 2 per cent over the previous year. Declines in smelter production in Europe and America were to some extent offset by increases in socialist countries. However, total

Australia			Europe[1]			Total		
Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption
thousand metric tons								
	0.5	1.2		480.0	479.5		665.6	673.1
148.1	5.2	3.6	344.6	429.7	522.2	1 064.8	876.0	885.6
149.6	41.9	12.4	346.8	520.0	598.4	1 292.8	1 105.4	1 106.8
171.7	65.1	27.3	456.7	648.0	772.2	1 541.0	1 332.9	1 337.4
189.9	79.3	39.5	451.5	544.5	703.6	1 928.9	1 711.3	1 674.5
263.2	102.6	68.9	966.0	1 220.3	1 295.4	2 952.1	2 702.2	2 527.4
392.3	196.0	99.5	1 387.0	1 859.5	2 013.9	4 340.9	4 184.0	4 129.6
509.9	242.9	114.6	1 656.5	2 266.6	2 393.5	5 205.5	5 165.6	4 984.3
487.2	257.5	113.9	1 674.8	2 316.7	2 409.0	5 341.3	5 077.3	4 912.8
448.2	255.5	116.0	1 736.5	2 285.6	2 425.8	5 429.6	4 967.1	3 008.0

Australia			Europe[1]			Total		
Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption	Mine production	Smelter production	Metal consumption
thousand metric tons								
5.7	5.7	0.7	4.9	19.5	56.6	102.5	104.6	104.0
5.8	4.3	1.5	4.9	28.6	52.8	128.2	125.9	122.3
2.8	2.9	1.4	2.7	43.1	60.4	150.3	151.8	147.1
3.0	2.7	2.1	3.0	56.3	71.9	158.5	162.1	156.0
2.5	2.5	2.8	6.9	49.6	50.5	144.3	145.1	135.2
2.0	1.9	3.2	12.9	75.1	81.5	185.8	190.3	169.4
5.0	3.5	4.6	26.8	68.2	105.1	211.4	214.4	223.5
8.3	4.2	4.2	29.9	72.6	111.9	228.6	232.7	239.9
8.8	5.2	4.2	30.0	70.4	114.7	236.4	238.9	236.6
9.5	6.3	4.3	31.3	65.9	115.9	241.5	240.4	241.4

world consumption for the year exceeded its production. Similar excess consumption over production was also recorded for lead and tin. The bulk of tin is produced by the developing countries but is consumed by the developed countries being 30 per cent in the United States and 12 per cent in the United Kingdom. Other larger consumers are the Federal Republic of Germany and Japan.

Mineral estimates and resources

In the developed countries, the *per capita* demand for copper is projected to triple by the year 2000. In the case of steel it is estimated that the total consumption of developing countries may rise from 31.7 million tons in 1965 to 83.9 million in 1980. Thus, faced with the most serious dilemma of population explosion, we are invariably concerned

with how much of these mineral resources are still available for future use. However, to estimate mineral reserves, both existing and potential, is not an easy task. The greatest difficulty results from the fact that they lie hidden in the earth and are not obtainable for accurate assessment as would be any other biological resource. In addition, measurements of limits of such materials require very expensive procedures. Analysis of ore deposits is further complicated by the factor of geologic time and by processes affecting their enrichment and dispersal. More uncertainty ensues when there is speculation as to whether technological advances will make present uneconomic deposits workable. This implies that the quantity of usable mineral resources is not fixed but it changes with innovation and progress in science and technology. Hence a need to review estimates periodically in the light of new developments. It may also be emphasized that most projections of international demand and supply are oriented towards requirements of the industrially advanced regions which provide the major markets. Developing countries are still busy exploring their resources; their consumption rates are insignificant compared to those of developed countries. Fig. 100 provides estimated recoverable reserves of selected metals at current prices and grades, while

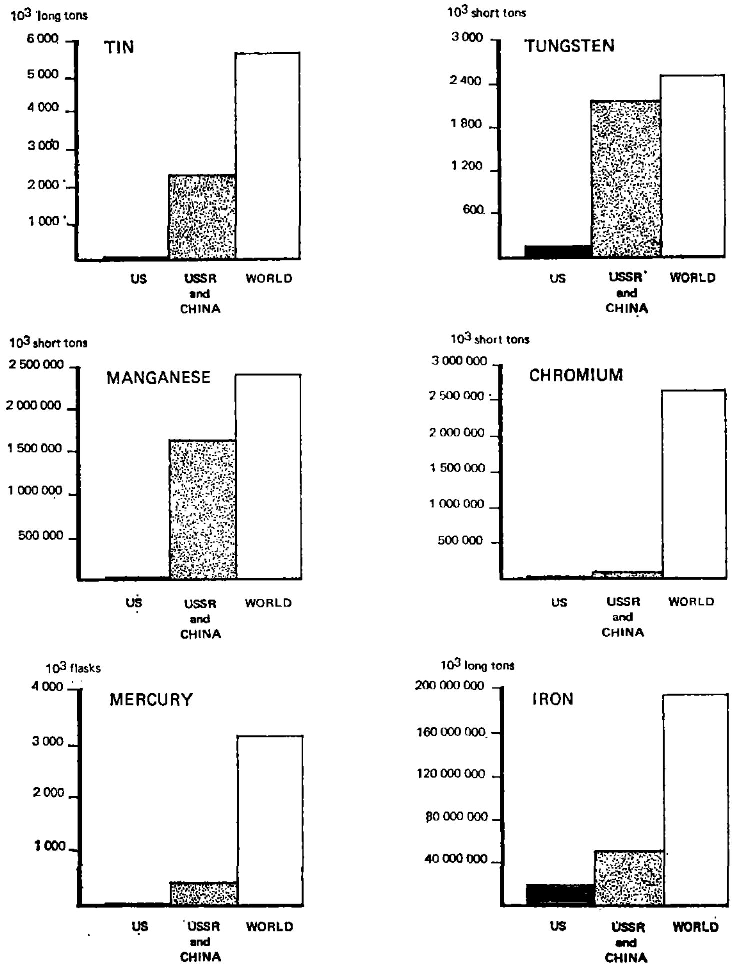


Fig. 100
Estimated recoverable reserves of selected metals at current prices and grades

Source: S. F. Singer (ed.). *Is There an Optimum Level of Population?* New York, McGraw-Hill, 1971, p. 20.

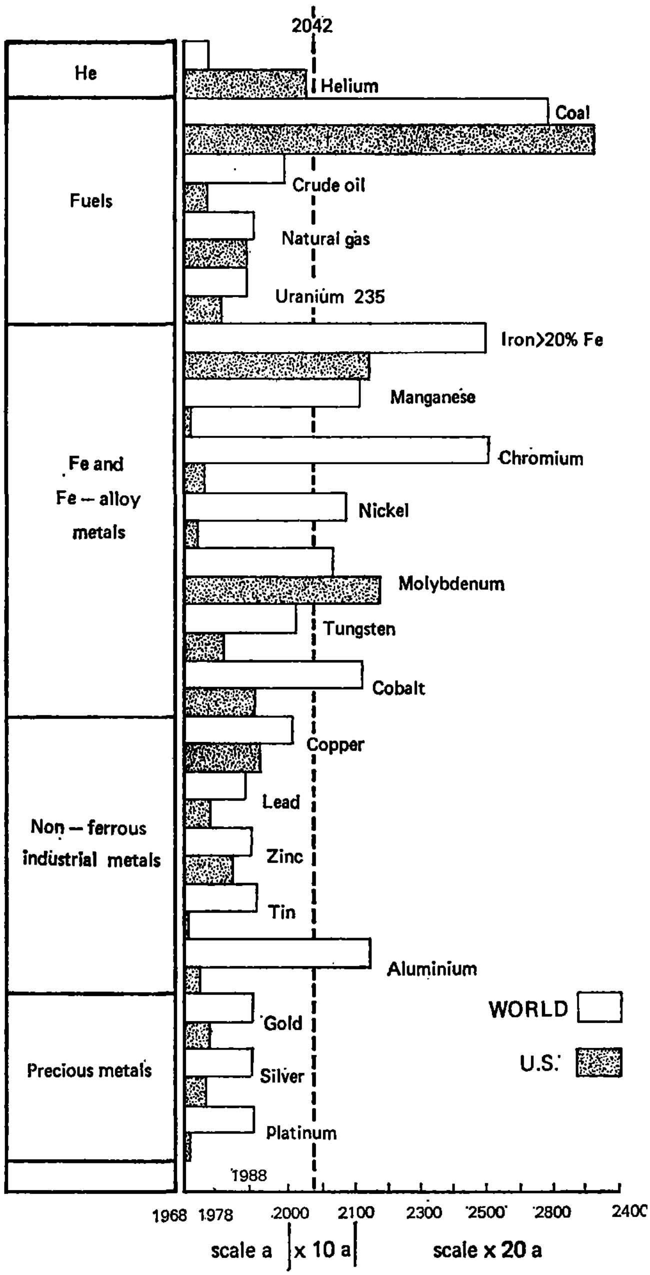


Fig. 101
Apparent lifetimes[1] of known recoverable reserves of twenty mineral commodities at current minable grades and rates of consumption

Source: S. F. Singer (ed.). *Is There an Optimum Level of Population?* New York, McGraw-Hill, 1971, p. 20.

1. Except for helium whose lifetime is estimated from U.S. Bureau of Mines data on reserves, conservation practices and expected increases in demand. Such lifetimes tend to increase with new discoveries and technological advances and to decrease with increasing populations and consumption rates; however, resources to left of vertical dashed line are in obvious danger of depletion).

Fig. 101 shows apparent lifetimes of known recoverable reserves of twenty mineral commodities at current minable grades and rates of consumption. It must, however, be emphasized that these lifetimes may increase with new finds, recycling, substitution and technological advances or may decrease with increasing population and consumption rates. For example, the United Nations Survey of World Iron Resources (1970) demonstrates a world total of more than 250,000 million tons of iron-ore reserves and over 500,000 million tons of potential ores compared with 85,000 million tons of reserves reported in the 1954 survey. The obvious reasons for this increase include new discoveries, inclusion of low-grade ore previously considered not economically feasible, and more information on known iron regions and inclusion of areas from which little or no information was previously available. Similarly, in the light of the conditions likely to prevail on our planet in several decades, production of non-ferrous metals would double in twelve years and increase eighteen fold in fifty years assuming an annual growth rate of 6 per cent. However, Fig. 101 indicates that, of the twenty minerals, only eleven are expected to last beyond the end of the century. Platinum, gold, zinc and lead are in obvious danger of depletion before the end of the present century. Today, coal, iron, copper, aluminium, petroleum and fertilizer minerals seem most basic to our civilization, and only a few industrial countries have adequate internal reserves for all these to meet their current consumption rates. None, however, seem to have reserves adequate for the next century.

Mineral depletion and future strategy

If events in Japan and India foreshadow those in other Asiatic nations, the exhaustion of reserves of currently commercial grades may come sooner than expected. According to many reports technology and vastly increased cheap nuclear power may stretch apparent lifetimes of such resources by allowing utilization of lower grades ores at cheaper costs both on land and in sea. However, such reports fail to take into account our deep concern about the environment. Uncontrolled exploitation for rising demands for minerals creates the problem of controlling the detrimental effects of mining on man's habitat. Pollutants from extraction and mining are progressively coming up against an ecological ceiling. Tighter pollution laws in some countries have already led to the shutting down of some plants. Use of nuclear energy for extraction calls for caution in terms of radioactive contamination and wastes. Evidently exponential growth both in mineral production and use cannot go on. A warning note to this effect was sounded as early as 1952 by the Paley Commission report long before the full impact of current population increase and environmental deterioration were realized.

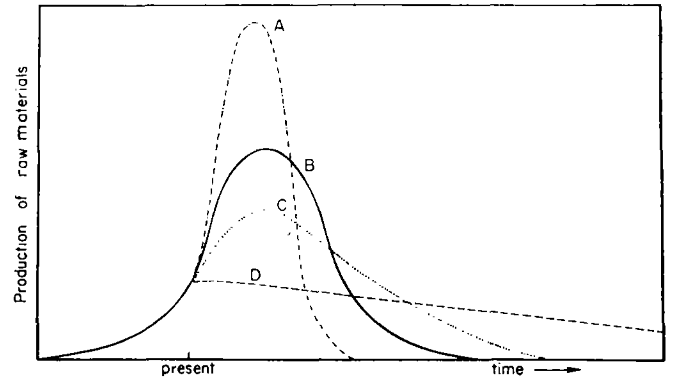
Cloud has recently inventoried supplies and reviewed prospects. To evaluate the current situation he suggests two concepts. The first one is the demographic quotient (Q) defined as

$$Q = \frac{\text{Total resources available}}{\text{Population density} \times \textit{per capita} \text{ consumption}}$$

As the quotient goes down so does the quality of life. The situation will continue to deteriorate as long as population and *per capita* consumption keep improving. Hence to keep Q constant we have the following choices: increase the death rate; increase resource availability; decrease birth rate; or decrease *per capita* rate of consumption. Whether we like it or not the third choice seems more convincing.

The second concept is the graphic model of depletion curves shown in Fig. 102. With the present procedures of mine use and waste, a huge boom and bust is projected as indicated by curve A. To quote E.P. Odum, 'the time-scale is uncertain because of

Fig. 102
Alternate depletion patterns for a non-renewable resource. (Efficient recycling combined with depletion curve D offers the best hope for the continuing availability of industrial metals on a long-term basis.)
Source: S. F. Singer (ed.). *Is There an Optimum Level of Population?* New York, McGraw-Hill, 1971, p. 12.



lack of data but the "bust" could begin within this century since certain key metals such as zinc, tin, lead, copper and others could be mined out in 20 years in so far as the readily exploitable reserves are concerned. However, if a programme of mineral conservation involving restrictions, substitutions and partial recycle were to be started now, the depletion curve could be flattened as represented by curve B. Efficient recycling combined with stringent conservation and reduction in *per capita* use especially by developed countries could prolong depletion for a long time as shown in curve C'.

Whatever precautions are made, depletion would still occur partly in manufacturing processes and partly through friction, rust, etc. Thus even with perfect re-cycling and conservation some millions of tons of metals would be required to replace the above-mentioned losses. The problems concerning the sea bed and its resources must be considered. Since there is growing evidence that oceans contain sufficient minerals to meet the world's needs and inevitable losses, endeavours should then be directed towards the development and improvement of technology to exploit these resources.

Minerals so far used by man have come from very near the earth's surface. Innovations in extraction technology should now aim at exploring the deeper layers of the earth's crust.

In addition, wasteful designs must also be discouraged. Skill in design and materials application should produce more durable and lasting products. All this will help extend the mineral resources base and consequently greatly flatten curve C shown in Fig. 102. Eventually, however, man must aim at striking a reasonable balance between population size and resource demands.

Further reading

ALFVEN, Hannes. Fission energy and other sources of energy. *Science and public affairs* January 1974, p. 4–8.
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE. *Energy and the future.* Washington, A.A.A.S., 1973.
ARMSTEAD, Cristopher H. (ed.), *Geothermal energy. Review of research and development.* Paris, Unesco, 1972, 186 p.
BHAVE, Deodhar *et al. You and your health.* New Delhi, National Book Trust.
BJÖRKMAN, O.; BERRY, J. High efficiency photosynthesis. *Scientific American* (New York), 299(4), October 1973, p. 80–93.
BORGSTROM, Georg. *World food resources.* London, Intertext, 1973.
BURINGH, P. *Introduction to the study of soils in tropical and subtropical regions.* Wageningen, Centre for Agricultural Publishing and Documentation, 1968.
BUTLEY, M. D. *Conserving soil.* New Delhi, Affiliated East-West Press Pvt., 1969.
CRESSEY, G. B. *Asia's lands and peoples.* New York, McGraw-Hill series in Geography, 1951.
DUVIGNEAUD, P. *Productivity of forest ecosystems.* Paris, Unesco, 1971, 707 p.
ECAFE. *Methods and problems of flood control in Asia and the Far East.* Bangkok, ECAFE, 1951.
——. *Helping Asia to harness its rivers.* Bangkok, ECAFE, 1967.
——. *Mining developments in Asia and the Far East.* Bangkok, ECAFE, 1970.
——. *Electric power in Asia and the Far East.* Bangkok, ECAFE, 1970.
——. *Proceedings of the 9th session of the regional conference on water resources development in Asia and the Far East (1970).* New York, U.N., 1971.
FAN SHUH CHING (ed.). *Farming in Singapore.* Hong Kong, Union Printing Co., 1969.
GANATHAN, V. S. *Economic geography of India.* New Delhi, National Book Trust, 1967.
GINDAL, I. A. *A new ecophysiological approach to forest-water relationships in arid climates.* The Hague, Junk, 1973, 142 p.
GOPALAN, C.; NARASINGARAO, B. S. *Dietary allowances for Indians.* Hyderabad, Nutrition Research Laboratories, Indian Council of Medical Research, 1971.
GOPALAN *et al. Diet atlas of India.* Idem, 1971.
HICKLING, C. F.; LANCASTER, Peter. *The seas and the oceans.* London, Blandford Press, 1973.
JELLIFFE, D. B. *Infant nutrition in the subtropics and tropics.* Geneva, WHO, 1968.
——. *Appréciation de l'état nutritionnel des populations.* Genève, Librairie Arnette (2, rue Casimir Delavigne, Paris 6), 1969.

JONES, O. A.; ENDEAN, B. (ed.). *Biology and geology of coral reefs.* New York, Academic Press, 2 vol., 1973–74.
KURIJAN, G. *India: a general survey.* New Delhi, National Book Trust, 1970.
LAL, SIDDAPPA and TANDON. *Preservation of fruits and vegetables.* New Delhi, Indian Council of Medical Research, 1967.
LEEDS, A.; VAYDA, A. P. (ed.). *Man, culture and animals.* Washington, American Association for the advancement of science, publ. No. 78, 1965.
LOWE MCCONNELL, R. H. (ed.). *Speciation in tropical environments.* London, Academic Press, 1969.
LOWENSTEIN, F. *Nutrition and infection.* Accra, WHO (P.O. Box 1628), 1967.
LUNDHOLM, Bengt (ed.). *Ecology and the less developed countries.* Stockholm, Swedish Natural Science Research Council, 1973, 133 p.
MALENBAUM, Wilfred. World resources for the year 2000. *Annals of the American Academy of Political and Social Science,* July 1973, p. 30–46.
MANI, M. S. *Ecology and biogeography in India.* The Hague, W. Junk, 1974, 773 p.
MAY, J. M. *et al.* (ed.). *Studies in disease ecology.* New York, Hefner, 1961, 613 p.
MISHRA, R. P. *Medical Geography of India.* New Delhi, National Book Trust, 1970.
MITCHELL, D. S.; THOMAS, P. A. *Ecology and water weeds in the neotropics.* Paris, Unesco, 1972, 50 p.
MOEN, Aaron, N. *Wildlife ecology.* San Francisco and Reading, W. H. Freeman & Co., 1973, 460 p.
MOORE, H. B. *Marine ecology.* 5th ed. London, 1966.
NATIONAL ACADEMY OF SCIENCES. *Soils of the humid tropics.* Washington, 1972, 219 p.
PEACH, W. N. The energy outlook for the 1980's. *Annals of the American Academy of Political and Social Science.* December 1973.
PEREIRA, H. C. *Land use and water resources in temperate and tropical climates.* London, Cambridge University Press, 1973, 246 p.
PIRIE, N. W. *Leaf protein: its agronomy, preparation, quality and use.* London, Blackwell Scientific Publications, 1971.
POEHLMAN, J. M.; BARTHAKUR, D. *Breeding Asian field crops.* New Delhi, Oxford and I.B.H. Publishing Co., 1969.
RAPAPORT, R. A. The flow of energy in an agricultural society. *Scientific American* (New York), September 1971.
RICHARDS, P. W. The tropical rain forest. *Scientific American* (New York), December 1973, p. 58–67.
SOEBJANI, Mohamad (ed.). *Tropical weeds, some problems, biology and control.* Bogor, Indonesia, Regional Center for tropical biology (BIOTROP), 1971.
SORRE, Max. *Les fondements biologiques de la géographie humaine. Essai d'une écologie de l'homme.* 4e éd. Paris, Armand Colin, 1971.
SUKHTME, P. V. *Feeding India's growing millions.* Bombay, Asia Publishing House, 1965.
TAMHANE, R. V., MOTIRAMANI, D. P., BALI, Y. P. and DONAHUE, R. L. *Soils: their chemistry and fertility in tropical Asia.* New Delhi, Prentice Hall, India, 1970.
UNESCO. *Problems of humid tropical regions.* Paris, Unesco, 1958, 102 p.
——. *Study of tropical vegetation.* Paris, Unesco, 1958, 226 p.
——. *Medicinal plants of the arid zones.* Paris, Unesco, 1960, 75 p.
——. *Tropical soils and vegetation.* Paris, Unesco, 1961, 115 p.
——. *Termites in the humid tropics.* Paris, Unesco, 1962, 259 p.
——. *The problems of the arid zone.* Paris, Unesco, 1962, 481 p.

UNESCO. *Environmental physiology and psychology in arid conditions.* Paris, Unesco, 1963, 345 p.; 1964, 400 p.
——. *Scientific problems of the humid tropical zone deltas and their implications.* Paris, Unesco, 1966, 422 p.
——. *Functioning of terrestrial ecosystems at the primary production level.* Paris, Unesco, 1968, 516 p.
——. Soil biology. *Review of Research.* Paris, Unesco, 1969, 240 p.
——. Pilot project for Biology Teaching in Africa. *Investigating living things.* (12 booklets for the pupils, junior level; one teacher's guide and one Microbiology teacher's study guide.) Paris, Unesco, 1969–70.
——. Projet pilote pour l'enseignement de la Biologie en Afrique.
La plante et le sol (les sols; l'eau et les plantes; nutrition minérale des plantes). Paris, AUDECAM, 1971, 84 p., 67 p. et 58 p.
Quelques aspects de la Biologie humaine (maladies; nutrition). Paris, AUDECAM, 1972, 103 p. et 119 p.
Les milieux naturels en Afrique intertropicale et à Madagascar (3 fascicules: écologie générale; écologie des savanes; écologie du milieu marin). Paris, AUDECAM, 1972–73, 57 p., 85 p. et 77 p.
——. *Use and conservation of the biosphere.* Paris, Unesco, 1970, 272 p.
——. *Soils and tropical weathering.* Paris, Unesco, 1971, 149 p.
——. *World water balance.* 3 vol. Paris, Unesco, 1972, 241 p., 296 p. and 169 p.
——. *International classification and mapping of vegetation.* Paris, Unesco, 1973, 93 p.
——. *Plant response to climatic factors.* Paris, Unesco, 1973, 574 p.
——. *Natural resources of humid tropical Asia.* Paris, Unesco, 1974.
——. Science and food for man. *Impact* (Unesco, Paris), vol. XXIV, No. 2, April-June 1974. This special issue contains numerous references.
——. *Aquatic vegetation and its use and control.* Paris, Unesco, in preparation.
UNITED NATIONS. *World energy supplies, 1968–1971.* New York, United Nations, 1973, 190 p.
VORSTER, D. J. M. (ed.). *Human biology of environmental change.* London, I.B.P. Central Office (7 Marylebone Road, London, N.W. 1), 1973.
WANG, G. W. (ed.). *Malaysia, a survey.* London, Pall Mall, 1964.
WHO. *Tropical food and nutrition.* Report No 20. Geneva, WHO, 1957.
——. *Report on intestinal infections.* Report No 288. Geneva, WHO, 1964, 37 p.
——. *Nutrition and infection.* Report No. 314. Geneva, WHO, 1965.
——. *Amoebiasis.* Geneva, WHO, 1969.
——. *Health hazards of the human environment.* Geneva, WHO, 1972, 387 p.
WHYTE, R. O. *Tropical grazing lands.* The Hague, W. Junk, 1974, 222 p.
WILLIAMS, C. N.; JOSEPH, K. T. *Climate, soil and crop production in the humid tropics.* Singapore, Oxford University Press, 1970.
X. The energy crisis: reality or myth. *Annals of the American Academy of Political and Social Science*, November 1973, p. 1–168.
X. Oil and the poor countries. Center for economics and social information. *Environment*, March 1974, p. 10–14.
ZEITZSCHEL, Bernt (ed.). *The biology of the Indian Ocean.* Berlin, Heidelberg, New York, Springer Verlag, 1973, 549 p.

Part III

Dynamics of human populations

Preamble

By 1840, the world's population stood at 1,000 million. It took less than a century to double it and merely thirty-six years to triple it. Likewise, it can be expected to double its present number by the end of the twentieth century. (See Fig. 103.) In most nations, despite careful planning and organization, much of their economic

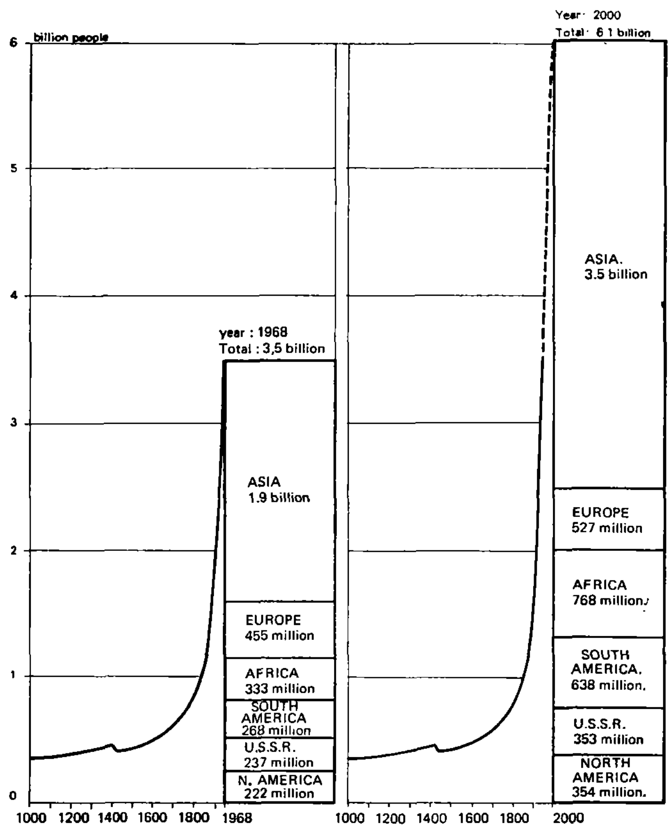


Fig. 103
World population growth.
In this diagram
one billion equals 1,000 million

growth is set back by the population growth, especially in the developing countries where it is two to three times as fast as that of the industrialized nations of the world. (See Tables 75 and 76.)

Table 75
Investment proportions required to keep *per capita* income at a constant level

Proportion of gross national product (GNP)[1] (in percentage)	Countries	Approximative annual population growth[2] (in percentage)
Over 10	Ghana, Morocco, Tunisia, India, Colombia, Brazil	3
7.5 to 10	Egypt, Iran, Malaysia, Thailand, Turkey, Mexico	2.5 to 3
5 to 7.5	Ethiopia, Nigeria, Sudan, Indonesia, Pakistan, South Korea, Chile	1.5 to 2.5
Less than 5	France, Sweden, United Kingdom, U.S.A., Federal Republic of Germany.	1 to 1.5

1. GNP estimates are for 1966 at 1964 prices.
2. Population growth figures are for 1967.

Table 76
Investments absorbed by population growth in developing countries

Countries	Amount (million U.S.$ at 1964 prices)
India	5 070
Brazil	2 060
Mexico	1 510
Colombia	720
Pakistan	680
Turkey	580
Nigeria	500
Philippines	380
Peru	340
Thailand	300
Malaysia	250
Morocco	250
Ghana	180
Sudan, Tunisia	90

13 Demographic parameters[1]

Introduction

Populations are collective groups, of the same species, living in particular areas, and having statistically unique characteristics. These characteristics include size, density, natality (birth rate), mortality (death rate), migration, sex ratio, age structure, growth form and internal distribution pattern. Uniquely for human populations, such social and economic traits as marital status, education, literacy, occupation and income can be added to the list. Human populations are distinct from the populations of other organisms in that the human endowment of intelligence provides the population with the capability for the development of culture. Culture enables a human population to consciously control its characteristics within biological and physical limits. Thus marriage laws influence birth rates and medical research influences death rates. Populations of plants and animals, on the other hand, are integrated with their environments and members cannot consciously make decisions that will change the characteristics of their population. However, activities of man can greatly influence the structure and pattern of these populations as well.

The attributes of populations can be described in quantitative terms and these measurements can be analysed to determine causality, and used to identify trends, forecast events, and aid in economic, political and educational planning. The science of gathering and analysing data on human populations is demography. Demographers describe the properties of populations as wholes and not the properties of individuals of the group. They may estimate with considerable accuracy the number of women of 42 years who will give birth to a son, or the number of men of 71 years who will die, in a given year, but they do not and cannot identify the particular individuals. Demographic data are only meaningful at the level of the population.

This chapter will focus on the structural characteristics (demographic parameters) of human populations and their measurements.

1. For the demographic parameters relevant to various countries, the reader may refer to national and international periodic publications (e.g. *U.N. Demographic Yearbook*; *Reports on Population/Family Planning*, *Country Profiles*, of the Population Council). (See Tables 78 and 79 in this chapter.)

Population size and density

The populations of most organisms are described in terms of numbers or mass per unit of space. Biologists may note a forest density of sixty cedar trees per hectare or a pond density of 3 million bacteria per millilitre, or sixty kilogrammes of fish per ha of water surface. Censuses of bees may be taken by counting the number of worker bees visiting a flowering tree per hour. The demographer counts people within political/geographical boundaries, generally countries. He attempts to assemble sufficient data to be able to describe populations in terms of age, occupation, ethnic and religious background, rural and urban residence. National figures are pooled to report regional, continental and world-wide data. Population densities of countries are computed by dividing the numbers of people by the area (in square kilometres) of the country. Human populations are not distributed uniformly within countries, and it is often instructive to consider the density of people in terms of livable space rather than total area. Large portions of national populations are often concentrated in urban areas where densities may be as high as 320,000 people per square kilometre (Hong Kong). In rural areas densities may vary with the productivity of the land. Depending upon the question asked, density may be computed and studied in relation to any of the variables that influence or reflect human behaviour, namely, climate, topography, economic activity, family structure and culture.

Nearly every country of the world tends to count its population on a regular basis. Methods and frequencies vary from country to country depending upon the nation's financial ability to set up a comprehensive census programme, the size of the country, and the degree to which its system of registration of births, deaths, marriages, migration, etc., are reported to a central data bank. The *de facto* system attempts to count all individuals who are present in the country at a given instant. In practice such censuses are massive efforts carried out during a single day. *De jure* censuses tabulate people by their place of permanent residence, and consequently require considerable effort and time to locate residents and decide which is a person's usual or legal residence. The advantage of the *de jure* scheme is that it collects data which are relatively unaffected by the temporary movements of people. Sample survey techniques are used in some countries and are effective when the statistical and research methodologies are carefully conceived. Registration systems which record births, deaths, and moves can be used to provide the essential data of the census. For the demographer, the ideal system is the complete register of vital events which in effect monitors every resident, from birth or immigration to death or emigration, recording every event of demographic significance in the life of that individual.

The usual types of censuses are subject to the obvious error that some people do not get counted. To correct undercounts, a post-census survey of a representative sample of the population is conducted to determine the proportion of people who were not counted during the census. Even in countries with well-developed census procedures the size of the undercount may be as low as 3 per cent, which for a nation with a population of 20 million people would amount to 600,000 individuals.

Intercensal growth rates

Censuses provide estimates of population size at points in time often ten years apart. Frequently it is desirable to have estimates of population size between censuses, or at points in time after the last census. One procedure for estimating intercensal population size, or project population size beyond the last census is to determine the intercensal growth rate. Let us assume that a country's population for 1960 was 21 million, and its 1970 census figure was 24.5 million. In the ten-year period the intercensal growth was therefore 3.5 million, and the population increased 0.35 million on the average each year. The average annual growth rate would be

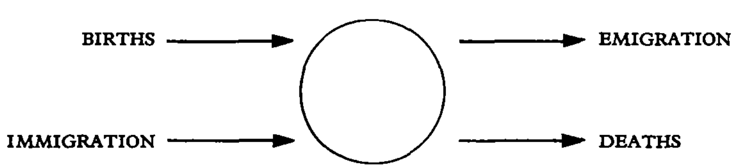
$$\frac{24.5-21.0}{(21.0)\ (10)}=0.017 \text{ or } 1.7\%$$

The growth rate can be used to estimate population sizes for years between censuses or for years beyond the last census. In the above example the 1971 population would be approximately 24.5+(24.5) (.017) or 24.9 million.

Projections of population size, using an annual average growth rate, would assume that rates are constant and unchanging. This is seldom if ever true and a growth rate may change suddently and unpredictably (especially in response to changes in infant mortality). A rate is valid only for the time it is measured. Calculations using age specific birth and death rates, and assuming the population is closed to migration, are considered more valid than average annual growth rates because they consider the age and sex distribution of the population along with the prevailing rates of birth and death.

The growth of populations is fundamentally the results of birth, death and migration. Changes in population size between time periods can be computed by adding the numbers of deaths and emigrations which occurred during the interval. The relationships can be expressed as an equation or as a diagram.

Population size (T_1)=Population size (T_0)+Births (T_0—T_1)—Deaths (T_0—T_1) +Immigration (T_0—T_1)—Emigrations (T_0—T_1)



Discussion of these variables and related characteristics occurs in sections following.

Fecundity

The biological potential or fecundity of a population is the maximum rate at which that population is physiologically capable of producing offspring. Maximum natality is a theoretical constant for populations and can be estimated by ascertaining reproductive limits. In human populations these limits reside with females since males are

capable of impregnating far more eggs during their years of virility (14 years of age to death) than the females of populations can produce, assuming natural sex ratios.

The fecundity of females, and thus of populations, is the number of children an average woman can bear in her reproductive lifetime. Assuming an average age of menarche of 15 years and an average age of menopause of 46 years, women can be assumed to have 31 fertile years. If women were to bear children every ten months for 31 years, each woman would produce 37 children, or 25 children if she waited six months between parturition and pregnancy. Despite the heroic performance of some superfecund women (the record is 39 single births) very few populations will approach an average of even seven children per female.

The restraints on fecundity are biological and cultural. Complete sterility or partial physical impairment of reproductive functions account for a significant measure of subfecundity in human populations. Recall that the incidence of barrenness among married couples is about 10 to 15 per cent. However, all human societies throughout history have imposed severe restrictions against the realization of biological reproductive potentials. There are taboos on the association of males and females, and taboos on sexual intercourse, not to mention a long tradition of birth-control practice. Despite these obstacles to fecundity, every society has also developed means by which it can encourage fertility. Societies reward reproduction by idealizing marriage, giving status and economic incentives to parenthood, and placing a high premium on virility. By manipulating obstacles and encouragements societies attempt to regulate their fertility at levels far below the theoretical limits. Fecundity is a useful concept in establishing the upper limits beneath which human performance can occur.

Crude birth rates

The simplest measure of actual fertility is 'crude birth rate'. It is calculated by dividing the total number of live births during a given year by the total male and female population as of 1 July of that year, and multiplying by 1,000 persons. In 1961 the mid-year population of Indonesia was estimated to be 96,371,421 and the number of babies born was 4,334,347. The crude birth rate was:

$$1\,000 \times \frac{4{,}334{,}347}{(96{,}371{,}421)\,(1)} = 44{\cdot}98, \text{ say, } 45.$$

Thus, for every 1,000 persons in Indonesia in 1961, forty-five babies were born.

The reason this measure is described as 'crude' is that it does not take into account the age and the sex structure of the population. Crude birth rates are easy to compute and to understand, but they have serious drawbacks. They do not reveal the rates at which women in the various age groups are having babies. For purposes of making projections of population growth, demographers are more interested in knowing the age structure of the female population and its current and potential reproductive performance, than in knowing how many births occurred per thousand people. Quite conceivably two populations could have identical crude birth rates but differences in

age structure would suggest entirely different demographic futures. Also, crude birth rates are unreliable for countries or areas where births are incompletely recorded. More sophisticated and useful measures of fertility are discussed in later sections of this chapter.

Mortality (death rate)

The physiological limit of longevity for humans appears to be about 120 years. If every individual of a population died at this old age, minimum mortality would prevail, and demography would be greatly simplified. But as in all naturally occurring populations, individuals are lost along the way in patterns that vary with the population and with the environmental conditions.

As with natality the simplest measure of mortality is the crude death rate, calculated by dividing the total number of deaths (after live birth) during a given year by the total male and female population as of 1 July of that year, and multiplying by 1,000 persons. Again using Indonesian data for 1961, there was a total of 1,770,552 deaths among the total mid-year population of 96,371,421. The crude death rate was

$$1\,000 \times \frac{(1{,}770{,}552)}{(96{,}371{,}421)} = 18.37, \text{ say, } 18$$

Thus, for every 1,000 persons in Indonesia in 1961, eighteen deaths occurred.

Crude death rate figures provide no clues as to who are dying within the population being described. Two populations might have identical crude death rates and this information may lead the reader to infer that the patterns of death in both populations are similar. However, the age and sex structures of the populations may be very different, with high infant and maternal mortality of the younger population equal to the old-age mortality of an older population. To describe mortality in more useful terms, age and sex specific death rates can be computed per 1,000 individuals of a given sex or age. Often ages are grouped into five-year intervals.

Life tables

A life table is a statistical device for picturing the complete life and death history of a population. It records the probabilities for survival, at specific ages, of males and females. One type of life table follows a cohort (all persons born during the same year) through to death of the last survivor. These generation life tables are not too practical since data would have to be collected for a period of about one century, and the environmental conditions which affect human mortality change so rapidly that the probability information derived from the data would have little application.

The preferred type of life table, the synthetic life table, is constructed of current age and sex specific death rates. It assumes that present conditions of mortality will prevail and that the hypothetical population described in the table will live and die in accordance with the life and death patterns of persons who lived and died during the

year from which data were taken for the construction of the table. Its shortcoming is that it projects one year of history for a lifetime.

Life tables are useful to insurance companies, pensioning foundations and others who must calculate premiums and annuities in accordance with life expectancies. Similar statistical treatments are applied to automobiles, buildings and machinery, all of which have life expectancies and for which replacement budgets must be established.

Survivorship curves

A survivorship curve is the result of plotting on a semi-logarithmic graph the numbers of survivors of a cohort against time expressed as the percentage of average life span. Such curves are useful in describing mortality in plant and animal populations under certain environmental conditions. Fig. 104 enables the comparison of the survival curve of the human with the survival curves of other organisms. The highly concave curve for oysters reflects extremely high mortality at earlier ages, somewhat duplicated by plants and most animals. The hydra displays a curve in which age-specific survivals remain constant. The curve for the human, like the curve for laboratory starved fruit flies, is convex and approaches the condition of organisms living out their inherited life spans. The shaded area above the human curve represents the human goal of achieving minimum mortality, by eliminating all causes of premature mortality, namely, disease, war, famine and accidents.

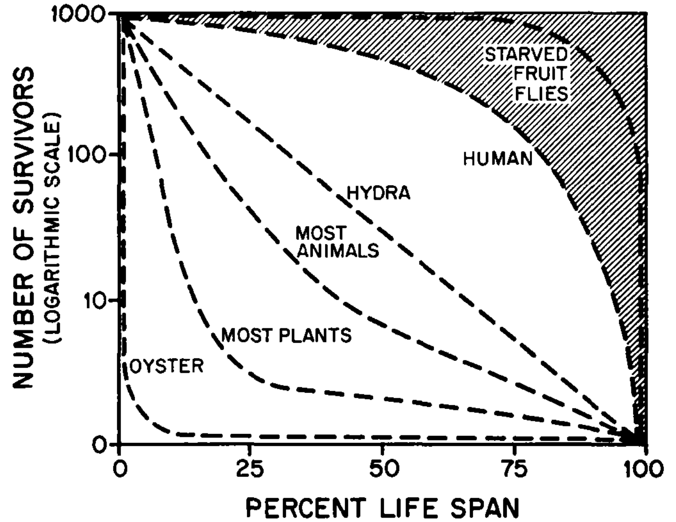


Fig. 104
Survivorship curve of human and other selected organisms
Source: Ralph Thomlinson. *Population Dynamics*. New York, Random House, 1965.

The chief causes of death have changed considerably within the last few decades. With effective control of such major killers as typhus, typhoid, famine, tuberculosis, influenza, pneumonia, cholera, smallpox, malaria, gastro-intestinal and other now controllable diseases, the leading causes for death among adults in developed countries are heart disease, cancer and cerebrovascular disease. Another focal point of high

mortality is the first year of life when hygienic care is critical to the survival of the infant. Infant mortality continues to be distressingly high among the poor and illiterate. (See Chapter 11 on 'diseases and environment', page 271.)

Major goals of medical research and public health for the immediate future will be the conquest of heart and vascular diseases, cancer, and the provision of improved sanitation and child care for the low socio-economic classes where infant mortality is high. Success in these efforts will certainly advance the human survivorship curve (Fig. 104) into the shaded area. And for those who search for ever greater challenges, there remains the question whether human beings must accept a limited longevity of 120 years.

Progress toward the achievement of minimum mortality has not been without attendant problems. By extending life for more people, without corresponding adjustments to fertility, human populations have increased in size to levels which threaten the quality of life, and, in the judgement of some, the very survival of human life. However, before engaging in a discussion of human population problems, additional variables and characteristics of populations should be examined.

Migration

In addition to births and deaths, human populations can be increased and decreased in size by the movements of people across international boundaries. Net migration is an annual rate of change expressed as the difference between the number of people who move in and settle in a country (immigration) and the number of people who leave a country to reside elsewhere (emigration).

NET MIGRATION = IMMIGRATION — EMIGRATION

The numerical value of net migration may be negative or positive.

Migration rates are difficult to compute since there has been little international agreement as to what constitutes permanent residential relocation, and over the years many countries have not established procedures for recording immigration and emigration. With recent curtailments of unrecorded migration, future measurements of net migration are likely to become accurate. However, it is now unlikely that such mass migratory currents as the flow of Europeans and the forced transfer of western Africans into North and South America during the nineteenth and twentieth centuries will ever be duplicated.

Fig. 105 illustrates the main currents of intercontinental migration since the beginning of the sixteenth century. Note also the large internal migratory movements which occurred in the U.S.S.R. and the United States.

Humans will probably continue to migrate in search of economic opportunity, more favourable climate, or political systems more compatible with their values. Recently, there have been some significant migrations: from northern Africa to parts of western Europe, from Europe, northern Africa and eastern Asia to Israel; between

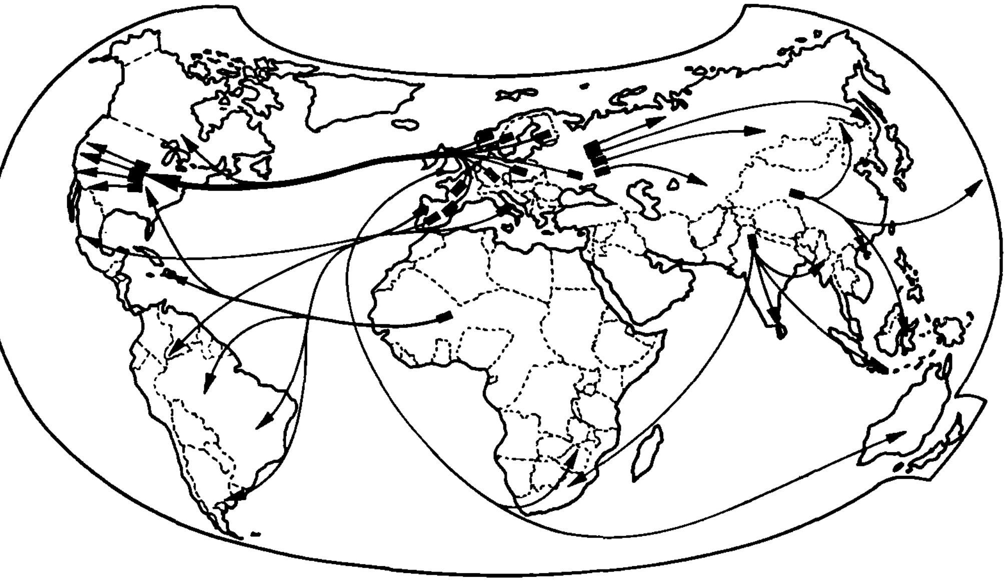


Fig. 105
The major currents of migration since the beginning of the sixteenth century
Source: Waytinsky and Waytinsky. *World Population and Production*. Twentieth Century Fund, 1953.

India and Pakistan; and into the United States from most countries of the world at a net migration of 400,000 a year. However, migration ceases to be a very important demographic variable. With so few regions of the world beckoning large inflows of migrants, only relatively slight future shifts in population size are anticipated to occur through international migration.

People also move wholly within their own countries, migrating from city to city, or between cities and rural areas. The causes of internal migration are often the same as those of international migration, but the flow is heavier because movements within countries are not as restricted by legal, linguistic and cultural obstacles as they are when persons leave and enter countries. In coastal western Africa the internal distribution of peoples has changed considerably since the turn of the century: by 1970, two-thirds of the population lived in coastal areas and one-third inland, against a reversed pattern in 1900.

Natural increase

Rates of natural increase are calculated by subtracting death rates from birth rates.

CRUDE RATE OF NATURAL INCREASE = CRUDE BIRTH RATE — CRUDE DEATH RATE

For Indonesia in 1961, the crude rate of natural increase was 45—18 or 27 per 1,000 population. This tells us that during 1961 the population of Indonesia increased by 2.7 per cent of its mid-year population.

The crude rates of birth and death do not provide the most adequate measures for natural increase. The age distribution of the population at the time calculations were made may not reflect the true natality and/or mortality of the population. If that population were to be closed to migration and became stable, i.e. achieve the state in which the proportion of persons in each age group became constant, then the rates of birth and death at each age would become constant, as would the rate of natural increase. Through the use of synthetic life tables it is possible to project a given population to the stable condition and compute rates of birth, death and natural increase for these fixed age groups. These computed rates are intrinsic rates and they only report what would happen if present age specific birth and death rates were to continue unchanged after the age distribution became stable. Intrinsic rates are syntheses of existing patterns of birth and death. They are not characteristic of future populations.

Growth rate

Calculation of the total annual growth rate of a population takes into account net migration. Growth rates, expressed as percentages, can be determined using the equation:

$$\text{GROWTH RATE} = \frac{(\text{BIRTH RATE} - \text{DEATH RATE}) + \dfrac{\text{NET MIGRATION}}{\text{POPULATION} \times 1{,}000}}{10}$$

Since birth and death rates are expressions of events per 1,000 persons of the population, net migration is divided by population size and multiplied by 1,000 to describe it in the same units as are used for births and deaths. The numerator is divided by 10 to express the growth rate as a percentage.

Age-sex distribution

The proportions of various age groups are especially important in determining the reproductive status of the population. Current and future birth rates and death rates of two populations of equal size may vary markedly when, for example, about 50 per cent of the members of one population is made up of children under the age of 15 while the other population includes an equally large proportion of individuals who are beyond their prime reproductive years. These populations will not only differ in measures of births and deaths, but also in their needs for economic, social and educational planning. See Table 77 for age distribution in the various regions of the world.

Three classes of age structure of populations can be usefully represented by polygons which show in their width and numbers or percentages of individuals and in their heights, the successive ages in the life span. Polygon A is the general form for

Table 77. Age distribution (percentage of total population) in various regions; 1970, 1985 and 2000

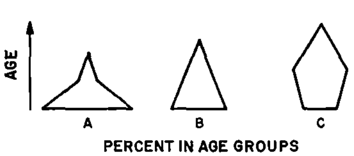
Region	1970				Assumed average 1985				Assumed average 2000			
	0–4 years	5–14 years	15–64 years	65 and over	0–4 years	5–14 years	15–64 years	65 and over	0–4 years	5–14 years	15–64 years	65 and over
World	14.0	23.0	57.8	5.2	13.4	22.9	58.2	5.5	11.4	21.4	61.0	6.1
Developed regions[1]	8.8	18.0	63.5	9.6	9.3	16.9	63.5	10.4	8.4	16.4	63.7	11.4
Less-developed regions	16.2	25.2	55.3	3.3	14.8	24.9	56.5	3.8	12.3	22.9	60.3	4.6
Eastern Asia	13.0	22.6	60.0	4.3	11.1	20.6	63.1	5.3	9.1	17.9	66.0	7.0
Southern Asia	17.5	25.8	53.8	3.0	15.5	26.4	54.6	3.4	12.0	23.3	60.5	4.2
Africa	17.8	25.9	53.5	2.8	18.1	26.9	52.1	3.0	16.1	27.0	53.7	3.2
Western Africa	18.2	26.3	53.2	2.4	18.8	27.0	51.6	2.6	16.9	27.7	52.5	2.9
Eastern Africa	17.8	26.0	53.4	2.8	18.3	26.8	51.9	2.9	16.9	27.5	52.5	3.1
Central Africa	17.0	24.9	55.1	3.1	18.1	26.0	52.7	3.2	17.0	27.3	52.4	3.3
Northern Africa	18.1	26.6	52.3	3.0	17.6	27.5	51.8	3.0	14.2	25.8	56.5	3.5
Southern Africa	16.3	23.9	56.0	3.7	16.6	25.4	53.9	4.1	15.0	25.7	55.0	4.3
Latin America[2]	16.4	26.0	53.8	3.8	15.8	25.6	54.4	4.1	14.4	24.7	56.6	4.4

1.	North America (1970)	: 30.0% (0–14 years),	60.8% (15–64 years) and	9.2% (65 years and over).
	Europe	: 24.3%	64.0%	11.7%
2.	Argentina	: 29.3%	63.4%	7.3%
	Bolivia	: 42.8%	54.2%	3.0%
	Brazil	: 42.0%	54.9%	3.1%
	Chile	: 39.4%	56.1%	4.5%
	Colombia	: 47.0%	50.4%	2.6%
	Ecuador	: 46.8%	50.4%	2.8%
	Mexico	: 45.8%	50.8%	3.4%
	Peru	: 45.0%	51.9%	3.1%
	Uruguay	: 28.2%	63.3%	8.5%
	Venezuela	: 45.1%	52.0%	2.9%

Source: A. C. Lee, L'accroissement démographique mondial et, en particulier, celui des pays en voie de développement. *Développement et Civilisations* (Paris), 47–48, 1972, p. 8–23.

an expanding population with a large proportion of young individuals. Polygon B represents the general form of a stable population experiencing constant age-specific birth and death rates over an extended period of time. If age-specific birth and death rates were equal, as in stationary populations, the area of the triangle would remain fixed. Polygon C is the general form of an aging population showing fewer youths to older age groups.

Fig. 106
General types of age pyramids
Source: Eugene P. Odum. *Fundamentals of Ecology*, 3rd ed., Philadelphia, W. B. Saunders Co., 1971.



With detailed age and sex information it is instructive to graph age-sex pyramids as appear in Fig. 107. As with the polygons above, the percentage of people represented by horizontal bars is plotted against age in five-year groups. Population pyramids further follow the convention of segretating the sexes, showing the age distribution of males on the left and females on the right.

The population pyramid for Ghana for 1960 (Fig. 107) reflects a youthful popu-

lation, with 9.5 per cent of the total population being composed of males under 5 years of age and 42 per cent of the total population of males and females under the age of 15. Comparable figures for Sweden for 1965 (Fig. 107), an older population, are 4 per cent and 21 per cent. Pyramid for France for 1973 and that for India, 1970 (Fig. 108) permit similar comparisons. That for Japan, 1958 (Fig. 108) shows disproportionately smaller male than female populations corresponding to Second World War years, and attributed to military losses. Bilateral dents in population pyramids can be

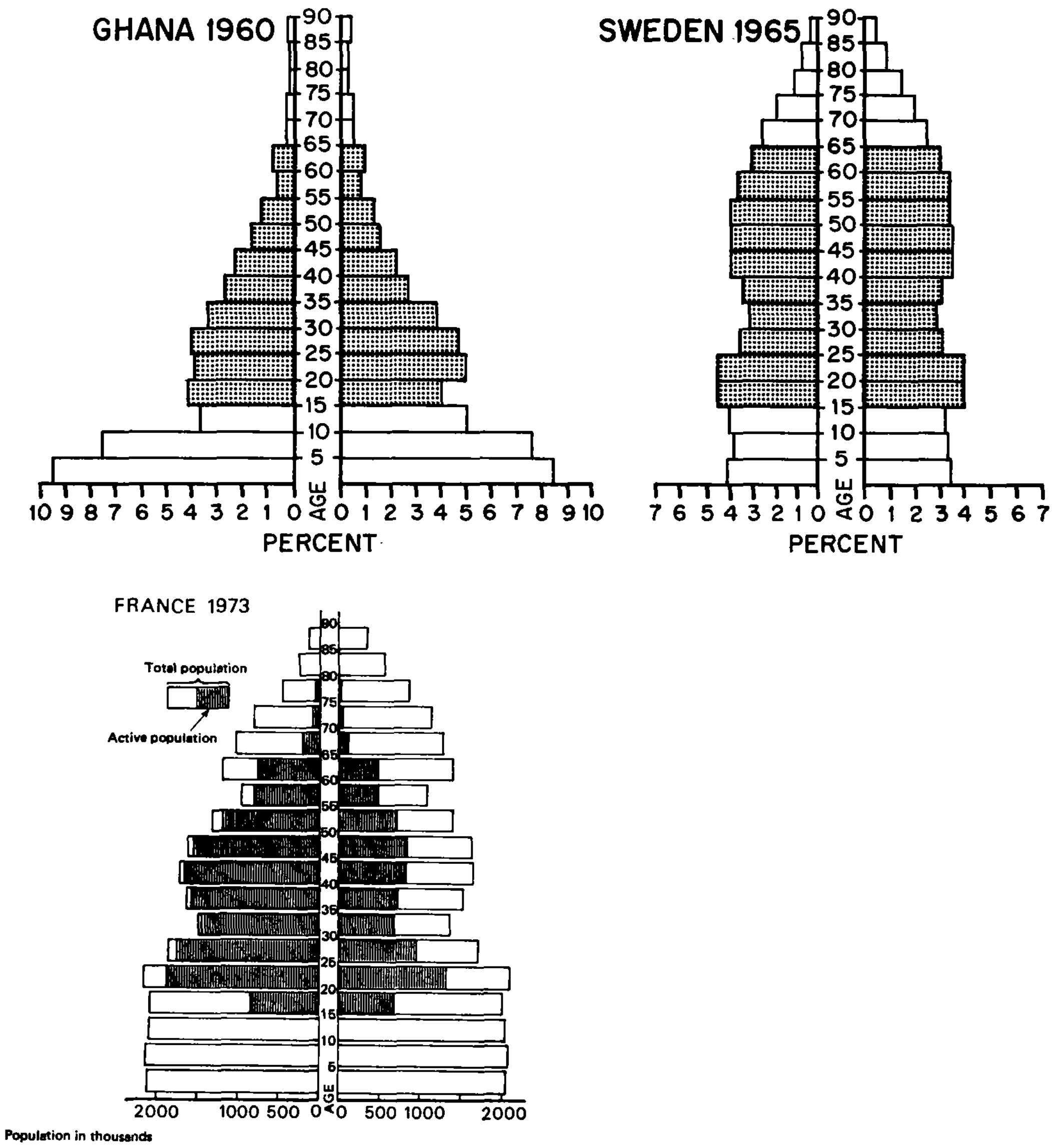


Fig. 107
Population pyramids

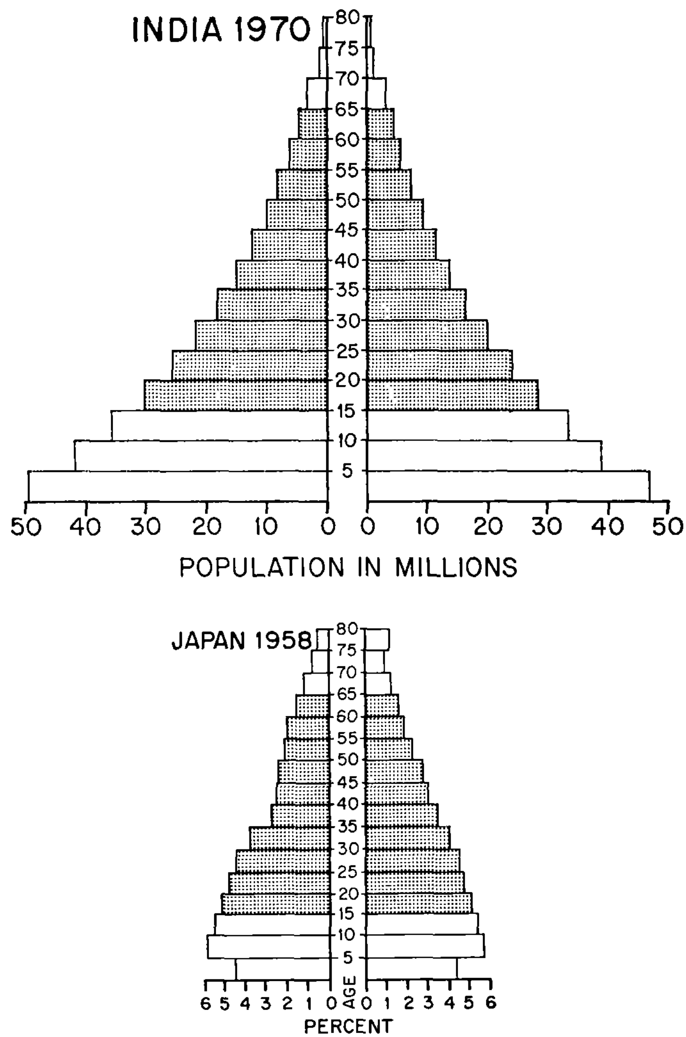


Fig. 108
Population pyramids

correlated with periods of economic depression, the war-time loss of potential parents and other historical events which temporarily reduce birth rates.

Population pyramids also illustrate the phenomenon of changing sex ratios with age. In Fig. 107 note that 9.5 per cent of the total population of Ghana in 1960 was males under 5 years of age, while the percentage of females under 5 years of age was only about 8.5 per cent. At the other end of the life span in the older ages the ratio shifts strongly in favour of females. The shift from numerical minority to majority for females can be observed in all other pyramids. At birth the sex ratio for humans (the number of males per 100 females) is about 105, whereupon it gradually declines due to higher male mortality, until the ninth decade of life when women often outnumber men by 2 to 1. The average sex ratio for populations is about 98, i.e. there are 98 males to 100 females. However, sex ratios can be distorted significantly by war,

migration and by varying the roles males and females play in the work economy of the country since one probable cause of earlier male mortality relates to the hazardous, exhaustive and mentally stressing roles males assume in most cultures.

Populations can be divided into three groups by age: the biologically and economically immature (younger than 15 years); the working and reproductive ages (15 to 64 years); and the post-reproductive and retired (65 years and older). Obviously, ages of entry into the work force and ages of retirement from the work force will vary among nations. These specific age groups were selected as international standards and as such facilitate comparative studies. The population pyramids of Figs. 107 and 108 are also shaded to illustrate the dimensions of these age groups and, for France, the active population against the total population.

Another derivative measure of age distribution in populations is the dependency ratio which compares the proportion of the population in the non-productive ages with the proportion in the working ages. Conventionally the working ages are considered to begin at age 15 and continue to age 65. The symbol ${}_{50}P_{15}$ reports the population size in the 50-year span beginning with the age of 15 years. (This portion of the population occupies the shaded portions of the population pyramids in Fig. 107.) Similarly, ${}_{15}P_0$ and ${}_0P_{65}$ report population sizes for the first 15 years and from age 65. Dependency ratios are calculated with the formula $\frac{{}_{15}P_0 + {}_0P_{65}}{{}_{50}P_{15}} \times 100$. Nations with high dependency ratios are burdened with relatively small work forces supporting large numbers of youth and/or the aged. Nations with high old age dependency gain relief by extending the years before work retirement, and nations with high youth dependency are often forced into lowering the age of entry into the labour force.

Age and sex distributions of populations not only affect the birth and death rates but have far reaching effects upon such variables as housing, *per capita* income, educational facilities, marriage, military manpower and even attitudes toward the young and the aged. For such reasons, age and sex distribution data are essential to national planning.

Fertility

Crude birth rates, describing the number of births per 1,000 population, are of limited value. They do not take into consideration the age and sex composition of the population. They report the number of births per 1,000 population including males and females of all ages. Males do not bear children nor do little girls and old women. To improve on the accuracy of the measure of the rate of birth, demographers developed procedures to measure the rates at which women in their child bearing years are producing babies. These are fertility rates and they enable demographers to scrutinize the patterns of reproduction and to project population replacement on a generational basis. There are several different measures of fertility each with its own advantages and disadvantages for projecting future population growth.

The general fertility rate is the number of live births of both sexes in a year to 1,000 women in the child-bearing 15 to 44 year age group. The general fertility

figure for Indonesia for 1961 was 195.58. If general fertility were to stay at this level for 30 years, the measure would correspond to 30 × 195.58 or 5,867 births per 1,000 women or 5.87 births per woman. Since general fertility measures assume constant fertility and fixed age structure (true only of stable populations), they are most useful in observing annual changes in current fertility.

The total fertility rate is the sum of all age-specific fertility rates current for that year. It measures the total number of children one woman would have had by the age of 45 years had she borne children for thirty years at the age specific rates which prevailed during that particular year. The total fertility rate is a generational measure while the general fertility rate describes the events of one year. In 1961 in Indonesia the total fertility rate was 5.47, meaning that if the population were to become stable at present rates, females would average 5.47 live births between the ages of 15 and 45. Total fertility rates, like life table projections, are hypothetical and are predictive only as long as current age specific rates continue.

Two additional indices of fertility are widely used by demographers to show more accurately than previously mentioned measures, generational fertility trends. Since males are of little demographic consequence the gross reproduction rate is the number of female children that a female just born may expect to bear during her reproductive lifetime assuming current age specific fertility will prevail and the possibility of her death before the end of her reproductive period is ignored. See Table 78 for gross reproduction rate and life expectancy in the various regions of the world. Net reproduction rate differs from gross reproduction rate in that prevailing mortality rates are computed to reduce the size of the cohort. Thus net reproduction rates are always smaller than gross reproduction rates. A net reproduction rate of 1.00 indicates that a population has achieved a stationary tendency. Indonesia's net reproductive rate indicates that the generational replacement rate for females is 1.88.

Table 78. Gross reproduction rate and life expectancy by region (1970–2000)

	Gross reproduction rate				Life expectancy for men and women (in years)			
	1970–75	1980–85	1985–90	1995–2000	1970–75	1980–85	1985–90	1995–2000
World	2.2	2.0	1.9	1.6	55.5	60.4	62.5	66.5
Developed regions	1.3	1.3	1.3	1.2	71.2	72.2	72.6	73.2
Developing regions	2.6	2.3	2.1	1.7	52.4	58.0	60.6	65.3
Eastern Asia	1.8	1.5	1.4	1.2	55.2	60.8	63.4	68.2
Southern Asia	3 0	2.5	2.2	1.6	51.8	57.8	60.6	65.8
Africa	3.1	3.1	3.0	2.5	45.9	51.2	53.7	58.5
Western Africa	3.2	3.2	3.1	2.7	41.8	46.8	49.3	54.3
Eastern Africa	3.1	3.1	3.0	2.7	44.9	50.0	52.5	57.5
Central Africa	2.9	3.0	3.0	2.7	41.8	46.9	49.4	54.4
Northern Africa	3.2	3.0	2.8	2.1	52.8	58.8	61.7	66.9
Southern Africa	2.7	2.8	2.7	2.3	50.4	55.1	57.4	61.9
Latin America	2.6	2.4	2.3	2.0	62.5	66.7	68.5	71.7

Source: A. C. Lee, L'accroissement démographique mondial et, en particulier, celui des pays en voie de développement. *Développement et Civilisations* (Paris), 47–48, 1972, p. 8–23.

Growth form

Populations of plants and animals exhibit characteristic patterns of growth which can be seen as curves resulting from arithmetic plots of density and time. The curves illustrated in Fig. 109 are greatly simplified and idealized. The growth forms of true populations rarely reflect the symmetry of the illustrated curves. A J-shaped growth form (Fig. 109) occurs when the density of individuals in the population, encountering little or no environmental resistance, increases exponentially to a level where the growth tendency is abruptly stopped. The upper level may be determined by the exhaustion

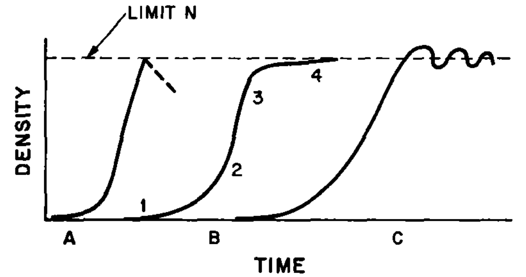


Fig. 109
Generalized curves of population growth form

Source: Eugene P. Odum. *Fundamentals of Ecology*, 3rd ed., Philadelphia, W. B. Saunders Co., 1971.

of a resource such as food or oxygen, or by frost, or merely by the end of the reproductive season. In most cases populations showing such forms do not maintain equilibrium at this highest density level, but decline as abruptly as they rise. Annual plants exhibits J-shaped growth form.

The second basic pattern of population growth is the 'S' or sigmoidal growth form, Fig. 109 (B). In this general pattern the population increases in density slowly at first (segment 1), then rapidly, often exponentially (segment 2) depending on the resistance of the environment. As resistance increases, the rate slows (segment 3) until, in the ideal situation, the density is maintained steadily (segment 4) at a level just beneath the upper asymptote, or the carrying capacity of the environment. Yeast and bacterial growth reflect this general pattern.

For many higher organisms, sigmoidal growth does not conform to Fig. 109 (B) but tends at first to become concave by the delay of the period of rapid growth due to logistic problems of mustering high fertility under low density conditions. Once rapid or exponential growth gets under way, there is a tendency to overshoot the upper asymptote. The process of adjusting natality and mortality to achieve and maintain a population level comfortably beneath the asymptote results in an oscillation which decreases in amplitude with time. Fig. 109 (C) shows a sigmoidal growth curve reflecting the phenomenon of delayed density effect and damped oscillation above and below the asymptote.

There are too few known facts about the growth patterns of human populations to facilitate the construction of growth curves similar to those which characterize the growth of populations of other species. However, numerous theoretical curves, describing historical, recent and future growth, have been proposed. Demography is

Table 79. Statistical data on the population of Asia (1970–72)

Region or country	Estimated population in the middle of 1972 (millions)[a]	Birth rate per 1,000 (1965–70)[b]	Death rate per 1,000 (1965–70)[b]	Infant mortality rate (death of under 1 year age for 1,000 live births)[b]	Annual growth rate (%) (1963–71)[b]	Percentage of population under 15 years[a,b]	Percentage of population in cities of 100,000 inhabitants and more[b]	Number of years for doubling of population[a]	Estimated population in 1985 (millions)[a]
Asia	2.154	38	15		2.3	40		30	2.874
South-West Asia	82	44	16		2.9	43		25	121
Saudi Arabia	8.2	1970 50	1970 23	—	2.7	—	1965 9.0	25	12.2
Bahrein	0.2	—	—	1968 35.5	2.9	—	—	23	0.3
Cyprus	0.6	1971 21.7	1971 6.4	1971 25.3	1.0	33	—	77	0.7
Iraq	10.4	1970 50	1970 25	1967 16.2	2.5	48	1970 40	21	16.7
Israel	3.0	1970 27.3	1970 7.0	1970 22.9	3.0	33	1969 34.0	29	4.0
Jordan	2.5	—	16.0	1970 22.1	3.4	47	1967 22.1	21	3.9
Kuweit	0.8	43.3	7.4	1970 39.4	9.8	38	1965 63.2	9	2.4
Lebanon	3.0	1971 26.5	1971 4.5	1960 13.6	2.9	—	1964 35.3	—	4.3
Oman	0.7	—	—	—	3.0	—	—	23	1.1
Qatar	0.1	—	—	—	5.0	—	—	23	0.1
Syria	6.6	47.5	15.3	1970 23.5	3.3	47	1970 31.0	21	10.5
Turkey	37.6	1970 40	1970 15	1967 153.0	2.5	42	1967 19	28	52.8
Gulf Emirates	0.2	—	—	—	3.0	32	—	23	0.2
Yemen	6.0	1970 50.0	1970 23	—	2.7	—	—	—	7.7
South Central Asia	806	44	17	—	2.8	43		27	1.137

India	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Iran	30.2	1971	42.4	1971	5.1		—	3.0	46	1970	23	25	45.0
Nepal	11.8	1970	45	1970	23		—	2.2	40	1967	1.3	32	15.8
Pakistan	146.6	1970	45–50	1970	14–18	1962–65	142	2.5–3 0	47	1970	10.5	21	224.2
Sikkim	0.2		—		—	1954	208.0	2.0	40		—	37	0.3
South-East Asia	304		44		16		—	2.9	44		—	25	434
Burma	29.1		40.3		17.4	1968	65.8	—	40	1958	5.3	30	39.2
Cambodia	7.6	1970	50	1970	25		—	2.5	44	1962	7	23	11.3
Indonesia	128.7	1970	40–45	1970	18–21	1962	125	2.0–2.5	44	1970	12	24	183.8
Laos	3.1	1971	45.9		17.2		—	2.4	—	1966	4.9	28	4.4
Malaysia	11.4	1970	33.8	1970	7.3	1970	40.8	2.4	42	1970	12	25	16.4
Philippines	40.8	1970	45	1970	11	1969	67.3	3.4	43	1971	14	21	64.0
Singapore	2.2	1971	22.8	1971	5.4	1971	19.7	2.2	39	1971	100.0	32	3.0
Thailand	38.6		42.8		10.4	1969	26.2	2.7	44	1971	8	21	57.7
North Vietnam	22.0		37.5		16.1		—	2.4	—	1960	6	—	28.2
South Vietnam	18.8	1970	35–42	1970	12–15	1960–61	42.8	2.6	—	1968	13	—	23.9
East Asia	962		31		14		—	1.8	35		—	41	1.182
China	786.1		33.1		15.3		—	1.8	—	1953	8.3	41	964.6
Taiwan	14.8	1970	27.2	1970	4.9		—	2.23	40	1970	35	30	19.4
North Korea	14.7		38.8		11.2		—	2.8	—	1960	6	25	20.7
South Korea	33.7	1970	30–32	1970	9–11		—	2.0–2.2	42	1970	30	35	45.9
Hong Kong	4.4	1971	19.0	1971	5.0	1971	18.4	2.1	38	1971	81	29	6.0
Japan	106.0	1971	19.2	1971	6.6	1971	12.4	1.1	24	1969	51.2	58	121.3
Mongolia	1.4		41.5		11.2		—	2.8	31	1962	19.4	23	2.0

Sources: a. Population Reference Bureau, Inc. 1775 Massachusetts Ave., N. W., Washington, D.C. 20036.
b. *United Nations Demographic Year Books*, 1970 and 1971. *Reports on Population/Family Planning*, No. 2, September 1972 (The Population Council, 245 Park Avenue, New York, New York 10017).

a young science, human generations are about thirty years in length, and data are insufficient to report accurately past and future events. The variables that establish the asymptotes for human populations are so numerous and complex that the task of identifying true carrying capacities of human environments are beyond the present capabilities of social and natural scientists.

Table 80. Distribution of the world population and annual growth rates in the various regions (1970, 1985 and 2000)

	Total population (millions)			Annual growth rate (in percentage)			
	1970	1985	2000	1970–75	1980–85	1985–90	1995–2000
World	3 632	4 933	6 494	2.0	2.0	1.9	1.7
Developed regions	1 090	1 275	1 454	1.0	1.1	0.9	0.8
Developing regions	2 542	3 658	5 040	2.5	2.4	2.3	2.0
Eastern Asia	930	1 182	1 424	1.7	1.5	1.4	1.1
Southern Asia	1 126	1 693	2 354	2.8	2.6	2.4	2.0
Africa	344	530	818	2.8	3.0	3.0	2.8
Western Africa	101	155	240	2.7	3.0	3.0	2.8
Eastern Africa	98	149	233	2.7	2.9	3.0	2.9
Central Africa	36	52	80	2.4	2.7	2.8	2.8
Northern Africa	87	140	214	3.2	3.2	3.1	2.6
Southern Africa	23	34	50	2.5	2.7	2.7	2.5
Latin America	283	435	652	2.9	2.8	2.8	2.6

Source: A. C. Lee, L'accroissement démographique mondial et, en particulier, celui des pays en voie de développement. *Développement et Civilisations* (Paris), 47–48, 1972, p. 8–23.

14 Population dynamics: growth and density trends

Introduction

This chapter examines the past, present and future of population change, and the theories which attempt to account for change. Of the demographic variables (natality, mortality and migration), natality emerges as the subject in need of greatest study, and discussion of determinants of fertility completes this coverage of population dynamics.

History of population growth

For hundreds of thousands of years the man was a hunter, a fisherman and a food gatherer. Even though he used tools, Old Stone Age man exercised little control over his environment, and the total world population probably never exceeded 10 million over a period of perhaps 2 million years. Then, about 10,000 years ago man learned how to grow his own food. He domesticated edible and usable plants and animals and established permanent settlements. With more food available he was undoubtedly encouraged and enabled to increase his numbers. As human civilizations advanced through the use of tools of stone, bronze and iron, agricultural technology improved and the world population increased fiftyfold to a size of approximately 500 million by the year 1650, the start of the industrial-scientific revolution. This rise in population, though gradual, was quite irregular, with ups and downs caused by years of abundance and years of famine, periodic wars and such epidemics as bubonic plague and cholera. Nevertheless, a jump in population size from 10 million to 500 million in just 10,000 years after 2 million years of essentially no growth has the elements of a population explosion. But a much greater increase was yet to come.

With the advent of the industrial-scientific revolution, people acquired considerably more control over their environment than was possible in a strictly agrarian system. World population estimates for the 350 years since industrialization reflect demographic consequences, as can be seen in the data tabulated below.

Year	Population (thousand million)	Year	Population (thousand million)
1650	0.5	1850	1.2
1700	0.6	1900	1.6
1750	0.7	1950	2.4
1800	0.9	1975	4.0

During the 170 years from 1650 to 1820, the world population doubled and reached 1,000 million. The next 1,000 million was reached 110 years later in 1930. The third was passed in 1965, after thirty-five years, and the next is almost certain to be achieved in 1975, only ten years later! (Refer to Fig. 103 and Chapter 13 on 'demographic parameters'.)

The times required for populations to double in size at given rates of growth are often shorter than appear possible. Population growth accelerates in the same way as money investments accumulate when interest is compounded. Thus, when a population grows at the relatively slow rate of 0.5 per cent, that population will double in 140 years. As the rates increase, doubling time decreases as can be seen below.

Growth rate (in percentage)	Doubling time (in years)	Growth rate (in percentage)	Doubling time (in years)
0.5	140	2.5	28
1.0	70	3.0	23
1.5	47	3.5	20
2.0	35	4.0	17

A graphical representation of human population growth through history, shown in Fig. 110 reveals the sudden surge in numbers of people particularly since 1950. At the current 2 per cent rate of natural increase (and even this rate is increasing) the world population grows at the rate of about 80 million each year, one million every 4.5 days, or 10,000 per hour! Unless there is an unexpected major reduction in fertility or a major increase in mortality, the population of the world will exceed 6,000 million by the year 2000. (Refer to tables in Chapter 13 on 'demographic parameters'.)

Even if the world's rate of natural increase were to decline suddenly to levels comparable with the rates of industrialized nations of the world today (say 0.5 per cent), the absolute increment in population would still be staggering (0.5 per cent applied

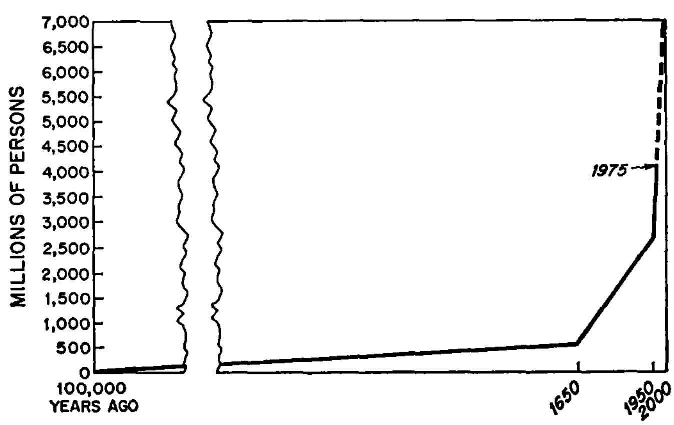


Fig. 110
A schematic representation of the increase in numbers of the human species
Source: Peterson. *Population*. New York, Macmillan, 1969.

to 6,000 million people adds 30 million people to the population each year). Obviously the upward rise of the curve cannot continue indefinitely. Sooner or later limiting factors of the environment—food, water, minerals, space, pollution—will almost certainly intervene to prevent further growth. But more hopefully human beings will, through intelligent planning, establish effective limits, based on the quality of life.

Causes of accelerated population growth

At first glance, the demographic revolution might be attributed to increased fertility. However, there is no evidence that the rates at which women have been having babies since 1650 have risen in any significant large section of the human population. In fact, the evidence suggests that there has actually been a decline in fertility. Certainly, migration cannot be responsible since on a world-wide basis all immigrations are also emigrations and total world population is unchanged. Therefore, mortality must be the demographic variable responsible for the spectacular growth of human population.

The scientific-industrial revolution contributed to the decline of mortality in developed countries in two significant ways. During the seventeenth, eighteenth and nineteenth centuries, agricultural production and trade improved. Gradually food, fuel and other previously scarce commodities became available regularly and in abundance. Standards of living rose, and more and more people were able to survive in environments that assured nutritious food throughout the year, fuel for the severe winter, water for the dry season, and mass transportation that enabled relocation in favourable distant lands. The more dramatic decline in mortality began in the late nineteenth century in the industrially developed nations when scientific medicine and sanitation made breakthroughs in the treatment and prevention of disease.

After it was discovered that the waterborne diseases, typhoid and cholera, were associated with the mixing of sewage and drinking water, communities took action to separate their sewage from their drinking water. The establishment of water control programmes in Europe and in North America clearly contributed to a decline in mortality and an accompanying increase in population size. Similarly, tuberculosis and typhus, deadly diseases associated with malnutrition, uncleanliness and poor housing, gradually ceased being major causes of death as higher standards of living became available.

By the twentieth century, tremendous advances were achieved in the control of nearly every communicable disease. Vaccines, antibiotics and insecticides, all inexpensively mass-produced and made available throughout the world in a period of just a few years, have suddenly and dramatically lowered mortality rates. Developing nations which up to the twentieth century had not industrialized and consequently were stabilized at high levels of fertility and mortality, characteristic of agrarian societies, were confronted with exploding populations due to rapid decline of mortality and the maintenance of near constant fertility.

South-East Asian countries provide a classic example of this effect. Prior to the end of the Second World War, malaria was probably the chief cause of sickness and

death in most tropical countries. In the years immediately following the war an extensive anti-malarial campaign to eradicate the *Anopheles* mosquito with the insecticide DDT was so successful in Sri Lanka that in a single year the mortality rate fell from 19.8 to 14.0 and life expectancy at birth rose from 43 years to 52 years. Over a period of 22 years following the war, the combined effects of anti-malarial programmes and modern medicine have reduced mortality rates in Sri Lanka from 22 to 7, raising the life expectancy at birth to the 1967 figure of 66 years.

In general, mortality rates in the developing areas of the world declined from 28 in the decade 1940–50 to 17 in the decade 1960–70. The startling impact of the means of disease prevention and control imported from developed countries can also be appreciated by comparing the pace of the same mortality changes in developed and developing countries. In Mexico in 1930, the life expectancy at birth was 33 years; 36 years later, in 1966, life expectancy had risen to 66 years, an increase of 27 years. In Sweden in 1780 the life expectancy was 36 years; one hundred and fifty years later, in 1930, the life expectancy had risen to 63 years. Thus it took Mexico only 36 years to accomplish what required 150 years of effort in Sweden.

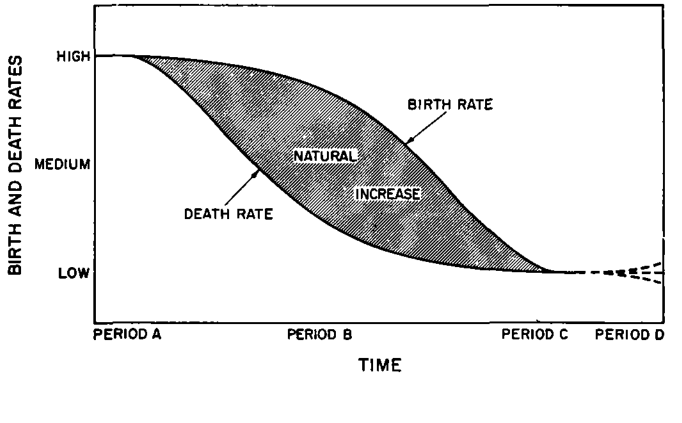
The decline in infant mortality had been especially significant demographically. By saving the lives of increasing numbers of children during their first year of life the population adds more individuals who are likely to survive through their reproductive years. Among five developing countries studied (Chile, Yugoslavia, Egypt, Costa Rica and Mauritius) infant mortality was reduced from 165 per 1,000 live births in 1932 to 83 in 1968. The potential for decline is well illustrated in Taiwan where between the same years infant mortality was reduced from 145 to 20. (Refer to tables in Chapter 13 on 'demographic parameters'.)

The impact of mortality decline during the scientific-industrial revolution can only be understood by noting what happened to fertility at the same time. Before 1650, world populations were relatively stationary. High mortality in pre-industrial societies was balanced by equally high fertility. People struggle against death by every known means. In order to survive they had to institutionalize patterns of reproductive behaviour which would lead to birth rates that at least equalled death rates. It thus became necessary for societies to establish pro-natalist policies. The position of the church, the rewards from government, the expectations from marriage, the thrusts of education were toward a level of fertility which for its time ensured generational replacement. When mortality declined, fertility also declined—but at a considerably lower rate. The reason appears obvious. Human value systems encourage death control and discourage birth control. Thus when opportunity to reduce mortality occurs society unhesitantly avails itself. But when fertility has to be lowered to corresponding levels the sluggish process of changing established attitudes has to be instituted and its result awaited with patience. Consequently fertility decline lags mortality decline, creating high rates of natural increase. When mortality declined slowly as in many European countries, fertility was never too far behind and population growth was moderate. However, when mortality rates plummet, as has recently happened in most developing countries, fertility rates are left far behind and rapid, sometimes explosive population growth results.

The demographic transition

The shift from a condition of high fertility and high mortality to a condition of low fertility and low mortality is termed the demographic transition. Fig. 111 is a simple diagram of the transition generalization. Specific rates and dates are omitted so that the diagram can be applied generally rather than to the events of specific nations. Period A is the old balance of agrarian non-industrial nations having equally high rates of mortality and fertility. Marriage customs and family size norms are deeply established to maintain high fertility. Severe mortality is maintained by malnutrition, starvation and endemic disease. Population growth fluctuates with such events as bumper crops, famines, epidemics and wars, but in the long run population size remains static. With high fertility, very large proportions of the population are in the reproductive age groups, a condition providing high growth potential—a spring awaiting release from the restraints of mortality. Some modern populations left Period A in the seventeenth century, many others are still there. Period B is the transition characterized by rapidly declining mortality rates and less rapidly declining fertility rates. The population expands in proportion to the lag of fertility behind mortality. Where death control measures are suddenly imported by populations in Period A, mortality drops precipitously, fertility stays behind at levels near the old balance, and population growth is very rapid. In the graph the shaded area between the curves depicts the gap between births and deaths and reflects the rate of natural increase. The geometry of this area (the so-called demographic gap) varies considerably among nations depending on such variables as rates of industrialization, availability of land to which population surpluses will migrate and the responsiveness of the institutions of the society to implement policies for lowering fertility. Period C represents a condition of zero population growth when mortality and fertility rates are equally low. It is a condition of maximum human efficiency and health. Early deaths are untimely tragic events. Families are small but populations are larger than they had ever been. Economics are efficient, industrial and stable, having produced the world's highest standards of living. Population pyramids evolve from broad based forms to narrow-based forms. Most of the nations of Europe are at present near the idealized balance of Period C. Period D merely acknowledges the truth that change is continuous,

Fig. 111
The demographic transition
Source: Thomlinson. *Population Dynamics*. New York, Random House, 1965.



that as cultures evolve the demographic variables of mortality and fertility also change, creating populations with new and different dimensions.

Each population experiences the demographic transition somewhat differently. In the absence of reliable demographic data prior to the twentieth century, accurate portraits of transitions cannot be drawn. However, Sweden, with data going back as far as the mid-eighteenth century, provided a model, illustrative of how one nation experiencing early industrialization progressed from the old balance to the new balance (Fig. 112). Note that throughout the transition, birth and death rates fell, roughly parallel, and growth rates over this 200-year period rarely exceeded 1 per cent in any given year. The gradual industrialization involved a complex series of social changes which operated gradually to reduce fertility. Typical of such European countries during these centuries of expanding population were the opportunities to emigrate to relatively unpopulated lands. Consequently, Sweden has proceeded from the old balance to the new balance while making only moderate gains in population and never having been burdened with numbers of people in excess of the carrying capacity of its economy.

Fig. 112
Pattern of change in birth and death rates in Sweden
Source: Population Reference Bureau Bulletin, April 1971.

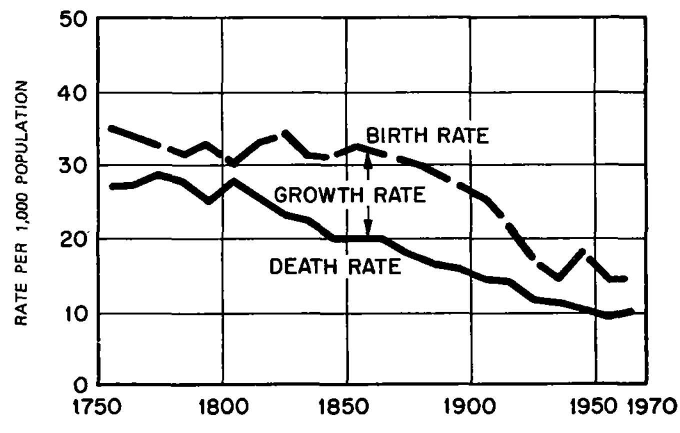
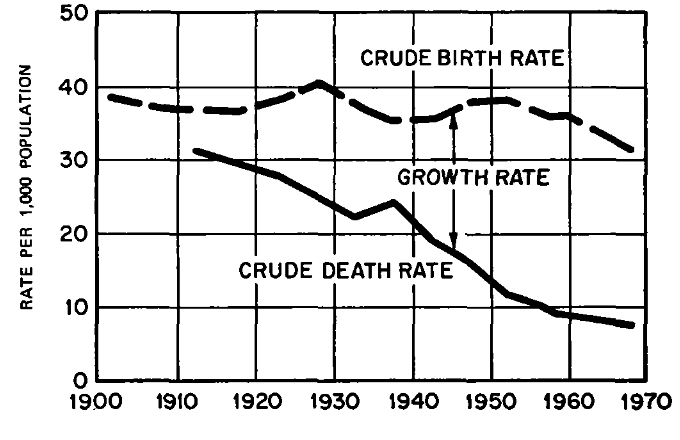


Fig. 113
Pattern of change in birth and death rates in Sri Lanka
Source: Population Reference Bureau Bulletin, April 1971.



In Sri Lanka (Fig. 113), the demographic transition did not begin until the twentieth century. In the seventy years covered by this graph, birth rates have declined slightly while death rates have fallen precipitously by comparison. The resulting growth rates of 2.5 per cent are typical of developing countries where rapid mortality decline occurs in the very early stages of industrialization before the occurrence of the social and economic adjustments necessary for the parallel decline of fertility. The resulting widening gaps between birth and death rates produce expanding populations without the wealth of technological development to support the population at standards of living comparable with industrial societies that are well along in the demographic transition. Furthermore, rapidly expanding populations are effectively locked within their national boundaries and have no new virgin lands to populate.

Populations which are in Periods A and B will likely eventually achieve the condition of low mortality and low fertility of Period C. The speed with which they pass through the transition will depend in large measure on their success in limiting fertility. Three alternative procedures to Period C appear possible: gradual technological development will affect fertility as it has in the developed countries; fertility will decline prior to technological development; and mortality will rise temporarily to relieve the pressures of burgeoning populations. The first alternative may be the most likely since this route has been travelled successfully by the nations now in Period C. However, the economic burdens of huge populations may destroy the incentives which are believed to contribute to the lowering of fertility. The second alternative is the goal of developing nations with population policies directed toward fertility reduction. The third alternative is a definite possibility.

Population theories

Modern theories which attempt to describe the causes of fertility have their origin in the ideas and beliefs of Thomas R. Malthus, who was the first prominent writer to express concern upon population growth. In a series of essays in the early nineteenth century, Malthus argued that populations tend to increase faster than food can be produced, and that the growth of populations must be checked by powerful forces. He believed that the human had an inherent reproductive drive from which exponential population growth results. Since food production can only increase at rates considerably less than the rate of natural increase, widespread famine would result as human populations bred to the exhaustion of their food supplies. He recognized the effectiveness of other checks which could intervene before starvation. These include disease, war, extreme poverty, and malnutrition. Malthus also recognized the potential of birth control in limiting population growth, but he was vigorously opposed to all forms of family limitation maintaining that unnatural and artificial release from sexual tensions would destroy societies both economically and morally. He felt that man was basically lazy and must be driven to work by the need to support a wife and children. The fewer children a man produced, the less his productivity. On the level of society he believed that population growth was essential to the maintenance of a viable economy.

Thus Malthus created a theory from which it was believed there could be no escape. The growth of human populations must exceed food productivity until or unless it is checked by powerful forces over which there are no technological, economic or morally acceptable controls.

Malthus's dire predictions have not yet materialized for reasons easily identified from the vantage point of the present. Advances in agricultural science have increased the area of arable land and the productivity of food through fertilizers, new high yielding breeds of plants and animals, insecticides and crop rotation. Industrialization has raised the standard of living and mitigated against the miseries which he projected would check the growth that has occurred. He did not foresee the mass migration from his native England and from other European countries, which reduced to a considerable extent the pressures at home to accommodate rapidly expanding populations. Finally, Malthus failed to recognize the possibility of widespread change in attitudes and practices with regard to birth control.

One hundred and fifty years later, the Malthusian theory is being re-examined with considerably more knowledge about the limitations of the environment and about how human populations behave reproductively. Some ecologists now believe that the consumption of the earth's energy resources and its polluting by-product may be more significant factors than food production in limiting population size. Nevertheless there are still compelling reasons to fear the failure of 'green revolutions' to maintain a level of subsistence in excess of a growing population's needs. Also it is now known that whereas populations of non-human organisms have high reproductive potentials, they do not necessarily exceed the carrying capacities of their environments but can be stabilized by environmental factors which become operative when densities achieve critical levels (certain species of aphids develop migrating winged offspring when densities reach specific levels). But with human population growth patterns, there is no evidence of any such density dependent control. In fact, there is some evidence that human populations grow faster as population density increases.

The application of parallels between non-human and human populations is seldom useful. Non-human populations (their fertility, mortality and migration) are regulated by external forces over which they have no conscious control. Human population in non-subsistence conditions (i.e. living below the carrying capacity of the environment), are largely controlled by the culture which the human himself creates. Human death is an inevitable biological event, but human culture has greatly modified patterns of survival by decreasing mortality and increasing life expectancy. The capacity to reproduce is biologically determined. Yet no human society reproduces to its biological potential. Biological factors determine the events of the menstrual cycle and the human gestation period, but human-made culture determines the age of marriage, the frequency of sexual intercourse, the number of children a couple will desire and the number of children the couple will produce.

The idea that human fertility is driven by irresistible biological forces to exceed the carrying capacity of the environment is no longer acceptable. However, there are those who believe that fertility rates, even though culturally determined, will always be higher than necessary to maintain populations of optimal size. Proponents of this

idea advocate that society must impose fertility control upon itself if the consequences of over-population are to be avoided.

An alternative contention argues that people will reproduce at rates which will guarantee the maintenance or the improvement of their standard of living. Thus in developed countries, gradual fertility rate declines accompanied by gradual industrialization with all its benefits for comfortable living. By the early 1920s, about ten such nations had achieved levels of replacement fertility. Developing countries with low rates of industrialization and high rates of natural increase pose the problem of expanding populations which inhibit their own economic development. Proponents of this 'standard of living theory' generally believe that low fertility behaviour must be motivated by the same kinds of incentives which have been operative in industrial nations throughout their demographic transition. That is, people of developing nations must have the knowledge and the assurance that low fertility leads to a better life. So motivated, and provided with the means of family planning, they will voluntarily aspire to small family norms and lower fertility rates. Implicit in this theory about fertility is the belief that coercive control cannot work in a society that is not intrinsically motivated to lower its fertility.

It should be instructive to examine the roles of social institutions in determining fertility and the forces which appear to effect changes in fertility under the influence of industrialization.

Sociology of fertility[1]

This section will examine aspects of the sociology of fertility characteristic of agrarian populations with high rates of mortality and fertility. Next, the cultural changes which affected the fertility of populations which gradually passed through the demographic transition to the condition of low rates of mortality and fertility will be described. Finally, discussion will centre on programmes designed to accelerate the decline in fertility in modern countries with high fertility and low, or potentially low, mortality.

Agrarian societies were characterized by marriages that were consummated at very early ages. This factor increased the number of years of exposure to pregnancy and contributed significantly to the high fertilization of the population. Young married couples were generally incorporated into the large socio-economic unit of the extended family and thus did not themselves experience the financial burdens of rearing children. Within the extended family high fertility was promoted through various pressures. Children contributed to the labour force of the family, and consequently children were economic assets—the larger the family, the greater its wealth. Mortality rates being high, the birth of many more children (especially males) than would survive to adulthood was essential to the later support of the older members of the family. The rank and esteem of women were often in proportion to the number of children they produced and in many societies their score cards recorded only the economically more

1. See also Chapter 19 on 'regulation of fertility' in Part V, page 413.

valuable sons. Sterility in women was grounds for divorce. Motherhood was the whole of women's role in society. Similarly, a major criterion of life effectiveness for the male was his ability to sire children, and manliness was gauged by the number of sons a man had produced. The unmarried state was a disgrace, and rarely could a man without a family assume prominent positions of leadership. Thus, in order for such agrarian societies to survive, their marriage practices, family organization, religious beliefs, laws, morality and education had to centre on the maintenance of high fertility.

As industrialization proceeded, mainly in European countries, the extended family gave way to the nuclear family unit in which a couple establishes a household that is physically and economically separate from the households of the parents of either spouse. This independence required the delay of marriage until the couple was able to assume the full responsibility for the rearing of children. Consequently the age of marriage began to rise, reducing the period of risk of pregnancy and quite likely contributing to the decline of fertility. The chances of child survival increased with declining mortality rates and in time couples realized that it was no longer necessary to have, say, five children when they really only wanted two or three. Economic opportunities provided by urban industries and the exploitation of the resources of the world made having large families far less rewarding than when the size of the family itself was a function of its economic worth. No longer did families look to their sons as their only source of old age security, since savings and pensions assured families of incomes after retirement or support in the event of disability. Other social reforms were gradually instituted to protect the rights of individuals against exploitation, and in effect encourage the formation of a small family. Child labour laws and compulsory education prevented parents from deriving economic benefit from their children. Usually the rearing of many children became a financial burden to parents, and the less children a couple supported, the more freedom they created for themselves to participate in activities that were rewarding alternatives to rearing (more) children. The unmarried state became no longer the disgrace it had once been. The present percentages of never-married women in Sweden and Ireland are 21 and 26, respectively, while in agrarian or subsistence societies the percentage probably has been less than 5. The much greater status given to women has been of considerable significance. No longer was there great emphasis on her role in reproduction and child care. Women had begun to join the labour force in nearly every occupation and profession.

The factors, described above, that have been acting in industrialized nations to lower fertility rates gradually to the level of mortality rates have not been operative in developing countries where mortality rates have rapidly fallen far below fertility rates. A relationship between small family size and socio-economic welfare is not obvious to most people, and for understandable reasons. The benefits of lowered fertility accumulate slowly and may not become tangible gains for individuals for perhaps generations. What can be done?

Obviously, developing nations will continue to attempt to accelerate their rates of industrial development. At the same time they will do all they can to employ medical and public health technology to minimize mortality and establish a healthy and vigorous populace. In recognition of the dilemma that very rapid population growth overtakes

the economic gains of increasing industrialization, some nations have established population policies and instituted programmes with objectives leading to the reduction of fertility rates.

Family planning programmes are deliberate efforts to reduce fertility without coercion by providing information and services to the people. Information includes the demographic facts and the interpretation that small families have social and economic advantages over large families; that while individuals may have as many children as they wish, whenever they wish, their own welfare, the welfare of society, and highest standards of living are best assured with small families. The services provided by family planning programmes include the inexpensive or free means of fertility control. These means include appropriate forms of birth control, and maternal and child care to lessen infant, child and maternal mortality and illness.

The effectiveness of family planning programmes is uncertain. Thus far, the thrusts of most programmes have been to bring modern means of birth control to the people. In developed countries which have achieved relatively high standards of living and low rates of mortality, but until recently have maintained high fertility rates (e.g. United States, Japan), the people have responded positively to the availability of birth control measures and have lowered their fertility rates dramatically. The significant difference between these people of well developed industrial nations and peoples of industrially developing nations is the much higher level of motivation to lower fertility in the more developed nation. In fact, there are many who argue that motivation is the most important variable in controlling fertility.

There is little evidence that family planning programmes have been effective in reducing fertility—China being a possible exception. In places such as Taiwan and Japan, governmental support to family planning followed rather than preceded the populations' recourse to the means of family size limitation. Studies of knowledge, attitudes and practices reveal that despite intensive campaigns to convince people in developing countries to avail themselves of the means for family planning, the desires for large families continue. Excepting some urban populations in developing nations, the people are not convinced that small families are to their advantage. Although the lack of motivation may in many instances be founded on ignorance, there remain some compelling reasons why couples do not wish to limit families to two or three children. In India, for example, where the need to produce at least one surviving son is deeply ingrained, a couple must produce on the average 6.3 children to assure a surviving son according to 1960 life tables. It follows that in order to generate the motivation necessary for significant declines in fertility, even more extensive reductions in mortality must occur and be maintained over a significant period of time. Whether the motivational response will occur before population size surpasses critical levels is the worrisome question. In some countries, the recognition of population education as an essential ingredient of the school curriculum may well induce much of the motivation necessary to effect fertility reduction.

Predicting the future

Estimating the future size, distribution and other characteristics of populations is an extremely difficult and hazardous task. To predict what changes will occur in populations is to predict also what changes will occur in the social, economic, political, technical and physical environmental conditions that are in turn determinants of fertility, mortality and migration. With so many inscrutable variables involved, the population forecaster must rely heavily on subjective judgements. Economists, for example, are not yet sufficiently in command of their discipline to predict with much accuracy future economic conditions more than several years in advance. Without such information as the long-term availability of investment capital and the future demands for labour, the demographer cannot incorporate these essential data into his projections of fertility.

The simplest type of population forecasting is the mathematical extrapolation of present figures, fitting the extended curve into a path that reflects past tendencies. The broken line beyond 1973 in Fig. 110 is such an extrapolation. The reader may extend the lines of birth and death rates for Sweden and Sri Lanka in Fig. 112 and 113 using whatever data and feelings he may have. Quite likely few predictions would agree, including those of demographers.

Other kinds of forecasts make use of age specific fertility and mortality data to calculate future population size. With these data it is possible to project prevailing conditions, or conditions to which the population aspires, to any future date. For example, if India achieved replacement fertility (a net reproduction rate of 1.00) by 1985, its present population of 585 million would reach 1,000 million by the year 2150. If present rates of fertility and mortality were to continue as they are at present, the population will be 1,500 million by the year 2000. The reasons for projecting inevitable growth, despite the assumption of the achievement of replacement fertility, relate to the peculiar age structures of a young population. The large proportions of youth yet to enter their reproductive years will produce greater numbers of children than will the population be reduced by deaths, even though the present youths might be producing less children than are required for replacement. Despite the many inscrutable variables involved, population projections are essential tools in planning for housing, education, agriculture and industrial productivity. But projections tend to become increasingly inaccurate and unreliable the further into the future they attempt to reach.

In general, three alternative models for the future appear to emerge.

In the first model, world population is seen to continue its geometric growth beyond the carrying capacity of the whole earth. At a critical load (one guess is 14,000 million people in the year 2050) the population suddenly and violently crashes to from one-fifth to one-half of its original size. Similar phenomena have been found to occur among such herbivorous mammals as rabbits and deer whose populations increased unmanageably when their predators were removed.

A second possibility predicts a safe arrival at Period C (see Fig. 111). For this to occur, it is assumed that fertility rates will decline and level out with mortality rates

producing a stationary world population that can be sustained by the resources of the earth. The lowest level and earliest date considered feasible, considering the world's present age structure, are about 7,000 million in 2025. An upper limit of 14,000 million in 2050 depends on technological breakthroughs such as the harnessing of thermonuclear power. However, few ecologists are comfortable with figures in excess of 7,000 million.

A third possibility envisages the populations of many developing countries overshooting the carrying capacities of their environments and suffering the trauma of population reduction through famine, disease, war and extreme poverty. In response to these shocking Malthusian checks, societies would react rapidly. The cultural traditions which maintained the population at high fertility levels would be abandoned and their transition would soon be achieved. The precedent for such a response is Ireland which, after its shocking potatoe famine of the 1840s, had its population reduced from 6.5 million to 3 million, which it has maintained by vastly restructuring its economy and those social customs which contributed to previously high fertility rates.

15 Environment and population patterns

Introduction

Only about 29 per cent of the surface area of the earth is land; the rest is sea. Furthermore, most of the land is relatively inhospitable to human habitation, so nearly 50 per cent of the human populations lives on about 5 per cent of the land area of the earth. The first portion of this chapter describes the world distribution of the human population.

Crude density pertains to the number of individuals per unit of total area. The quantitative entities of density are thus population size and the amount of space. Density can also be expressed in terms of habitat space, which is usually termed ecological or specific density. In view of recent rural to high density urban population trends and because of shelter as a basic human need, the next section of this chapter describes space and housing.

The third section describes how temperature, water, food production and land and soil influence the distribution of human population density. In turn, the influence of population density on these factors is considered since they result in changes which affect human populations.

The last two sections consider some physiological, behavioural, psychological, and social conditions which arise in high density situations.

World distribution of human population

If people were distributed evenly throughout the land surface of the earth, their density would be a little over 19 per km^2. However, major land regions such as the Antarctic continent and Sahara Desert are almost uninhabitable and other regions are densely populated. Table 81 presents population densities of some selected countries. In Argentina, the density is similar to that of Brazil, and Australia is even lower. The over-all density in Europe (excluding the Soviet Union) is around 82 per km^2, while it is around 150 to 185 per km^2 in much of the Orient. (Refer also to Tables 78-80 in Chapter 13 on 'demographic parameters', pages 322-326.)

Table 81
Population densities of selected countries

Country	Density per km^2
Brazil	12
Germany	185
United Kingdom	222
Japan	222
Puerto Rico	250
Belgium	285
Indonesia	420

In a pre-agricultural hunting and food-gathering society, about 5.2km² of fertile land were required to support one individual. Agriculture and, more recently, technology have permitted much greater human population densities.

The change toward a large proportion of a population living in cities is termed urbanization. Societies which are economically and technologically developed tend also to be highly urbanized.

However, urbanization is currently an extensive and widespread process, not necessarily corresponding to the degree of technological development. The proportion of the world population in urban areas of more than 100,000 changed from 2 per cent in 1800 to around 18 per cent in 1960. In Egypt, the urban population has grown twice as fast as the national population during the twentieth century. The average

Table 82 Evaluation of the urban population and of its annual growth rate (%) in the various regions of the world (1970–2000)

	Urban population (millions)			Percentage			Annual growth rate (%)	
	1970	1985	2000	1970	1985	2000	1970–85	1985–2000
World	1 352	2 169	3 329	37.2	43.8	51.1	3.3	2.9
Developed regions	717	943	1 174	65.7	73.9	80.7	1.8	1.5
Developing regions	635	1 226	2 155	25.0	33.4	42.6	4.5	3.8
Eastern Asia	266	461	722	28.6	39.0	50.7	3.7	3.0
Southern Asia	238	458	793	21.2	27.1	33.7	4.4	3.7
Africa	77	160	320	22.2	30.1	39.2	5.0	4.7
Western Africa	20	45	96	20.0	29.0	39.9	5.4	5.2
Eastern Africa	9	22	50	9.6	14.5	21.2	5.7	5.7
Central Africa	6	14	33	16.5	26.9	40.7	6.0	5.8
Northern Africa	30	62	113	35.4	44.0	52.9	4.8	6.1
Southern Africa	10	17	29	45.7	51.9	57.9	3.5	3.5
Latin America[1]	158	291	495	55.9	66.9	75.9	4.1	2.6

. In 1970, Bolivia (35.5%), Brazil (47.6%), Cuba (53.4%), Mexico (62.3%), Peru (49.2%), Panama (50.2%), Uruguay (79.9%).

Source: A. C. Lee, L'accroissement démographique mondial et, en particulier, celui des pays en voie de développement. *Développement et Civilisations* (Paris), 47–48, 1972, p. 8–23.

annual growth of the two largest cities in Cameroon has been 6 per cent in recent years. In Brazil, the proportion of the national population in cities over 100,000 changed from 8.7 per cent in 1920 to 18.6 per cent in 1960.

Table 82 gives the evaluation of urban population and of its annual growth rate expressed as a percentage in the various regions of the world, for 1970, 1985 and 2000.

Limiting factors

Human population density and distribution are related to such factors as past growth rates and distribution patterns, transportation advantages, political policies, technology and various environmental aspects. As the human population density of an area increases, some environmental factors which initially facilitated human habitation often become limiting. For example, a town may develop at a particular place because of the availability of freshwater from a spring. If the town continues to grow and additional sources of water are not found, inadequate water eventually becomes a factor limiting further increase in population density in the immediate area. When the effect of a factor depends upon the population density, the action is referred to as 'density dependent'. In other cases, factors limit or restrict human population distribution, but their action is independent of the density of the population. For exemple, the effects of high altitude on the human body restrict human habitation to areas below 5,300 m and this effect is not dependent on density. Such actions are called 'density independent'.

In regions of urbanization, where human population densities are high, massive quantities of water for drinking, cleanliness and waste removal are required. Furthermore, the high population density does not permit natural processes of purification to take place between points at which wastes are released into the water and points at which water is used for drinking. As a result, extensive systems of sewage treatment and water treatment are necessary if diseases due to water impurities are to be avoided. In many cases, such systems have not been available but urbanization has gone ahead. About ten years ago, only about 33 per cent of the urban residents of seventy-five countries of Asia, Africa and Latin America had piped water in their homes. Another 25 per cent had water outlets within a reasonable distance, but over-all, for 89 per cent of the people, in these countries, the water quality was unsatisfactory for safe human consumption. In Latin America, 85 per cent of the rural population and 39 per cent of the urban population lack water services.

Technological advances bring additional water supply problems. Industries, food-processing plants and domestic uses all increase water supply demands and add waste materials to the water. Water then may be regarded as a density dependent factor in regard to human populations.

Poor agricultural soils generally do not support very dense human populations. In arid and semi-arid climates lack of water clearly limits productivity and population density. Such climates cover about 33 per cent of all the land and have about 6 per cent of the total population. However, the extensive lateritic soils of the low latitudes also

have not been highly productive in the past. The year-around wet and very warm regions of the world constitute about 10 per cent of the land area and 2 per cent of the world human population. Approximately an additional 20 per cent of the land occurs in polar and subpolar regions and is unsuited for agriculture. These regions contain less than 1 per cent of the world population. The remaining lands, making up about one-third of the world's land surface, contain about 90 per cent of the world's population. These are largely the seasonally wet subtropics and the temperate lands.

Thus, soils might be viewed as density dependent mainly in the indirect sense of limiting food production, which in turn affects high-density human populations more severely than low-density population, if other conditions are the same.

Human settlements

Pastoral nomads today usually lead a semi-nomadic life. In South-East Asia, Tibet and northern Africa, pastoral nomadism is dependent upon settled agriculture. The nomads procure food grains and other goods from the settled population and in turn supply the latter with livestock products. In the present century, the Soviet Union has taken measures to help them become settled as stock-breeders. In Iran during the first quarter of this century there has been a decline of the nomadic population from one-third to one-fifth of its former size. Nomads are still found in considerable strength in Chinese Sinkiang.

The efforts on the part of the Australian Government to settle the aborigines have had only limited success. Many of them have continued to prefer a wandering life. Similarly, members of a Rajasthani hill tribe in India who began a wandering life by following the footsteps of the patriotic and freedom-loving Maharana Pratap during the days of Akbar are still reluctant to lead a settled life, even though they are provided with free housing and other facilities by the State Government.

Availability of certain environmental facilities resulted in the development of permanent human settlements, durable and elaborated houses and establishment of cities; they provided the basis for some of the world's oldest civilizations. Ample arable land, adequate water supply, and food transport by water are examples of environmental contributions to such civilizations as that of the Nile Valley of Egypt, the Mesopotamian civilization of the Tigris and Euphrates rivers, and the Indus Valley civilization.

Houses, monuments and temples during this period were constructed of a wide array of materials. Among them were stones cut into geometric shapes; bricks, tiles and terra-cotta made of moulded and burnt clay; wood; plaster made of lime, gypsum, sand and sometimes cement; metals. The major metals were copper, tin, bronze, lead, zinc, brass and iron and its derivatives. The contribution of stone is especially prominent. It entered into the construction of the pyramids of Egypt, and was used in the Greek architecture utilizing white marble. The Indian cave temples of Ajanta and Ellora were carved out of solid rock. Brick dominated the buildings of Mesopotamia, Harapa and Mohenjodaro. Wood was used extensively in Scandinavia, Japan, Thailand and

America. Metals were used mostly for doors, windows, railings, hardwares and fixtures. Romans used glass for windows.

The industrial revolution led to the rapid growth of cities, new modes of transport and commerce. It also brought an increased migration of people from rural to urban areas. Europeans migrated great distances in large numbers and established new human settlements in virgin areas of the globe such as North America, Australia and New Zealand.

The World Health Organization expert committee on the public health aspects of housing, at its meeting in Geneva in 1961, defined 'housing' as 'residential environment' meaning 'the physical structure that man uses for shelter and the environs of that structure including all necessary services, facilities, equipment and devices needed and desired for the physical and mental health and social well-being of the family and the individual'. This concept thus makes 'housing' the core element of not only family, community and national planning but also calls for regional and international development programmes.

In highly urbanized countries such as the United Kingdom, Australia, Venezuela, the United States, Canada and Japan, approximately 70 to 80 per cent of the total population live in urban areas. In these areas, 1 to 1.5 room(s) is (are) available to each person except in Japan where the figure is nine-tenths of a room. There is little difference in the room space accommodation between the rural and urban population of these countries. In the Soviet Union, three-quarters of a room is available for each member of the total population.

In Chile, Israel and Iran, between one-half and three-quarters of a room is available per individual. Comparable data are not available for China but, in India, one room is shared by three persons.

The average number of persons per household in the ECAFE region is on the decline and, as a result, housing requirements in the area are expected to grow faster than the population. In a recent report (the United Nations Asian Population Studies Series No. 11, 1972) an over-all increase of 31.8 per cent by 1980 over 1970 in the three high fertility sub-regions of ECAFE has been projected. In Japan, however, the expected increase is even larger (33.2 per cent) in spite of its slower rate of population growth. The housing demands of many high fertility ECAFE countries become more enormous if improvements to the existing housing are also taken into account. With the possible exceptions of Hong Kong, Malaysia and Singapore, the present situation in these countries is generally deteriorating, especially in the urban centres where pressures of rural-urban migration are mounting high.

In India at the beginning of third five-year plan (1961), there were 14.1 million houses in urban areas out of which only 6.3 million were good houses. In rural areas the figures were 65.1 and 12.2 million respectively, giving a total number of 18.5 million good houses compared with 84.5 million households in the country. This indicates a shortage of 66 million houses, 9.3 million in urban areas and 56.7 million in rural areas. The bulk of the standing inferior houses in rural areas were mud huts. A large section of industrial workers in the cities also lived in sub-standard accommodations. The shortage of housing at the beginning of the fourth plan (1969) was

estimated roughly at 83.7 million units, of which 11.9 million were in urban areas and 71.8 million in rural areas. The over-all housing shortage was estimated to increase by more than 2 million units, annually.

Table 83 shows the extent of electrification in towns and villages in India, based on population. Obviously it cannot be inferred that each home in an electrified town has electricity, but these data provide some indication of the optimum living standard available.

Table 83
Electrified villages and towns in India

Population range	Total as per 1961 census	Number of towns and villages electrified by			
		1961	March 1969	March 1970	March 1971
Up to 9,999	567 217	25 661	71 812	90 383	105 345
10,000–49,999	2 114	1 785	2 029	2 036	2 041
50,000–100,000	139	138	139	139	139
Over 100,000	107	106	107	107	107

Source: India: a Reference Annual, 1973. New Delhi, Ministry of Information and Broadcasting, Government of India, p. 252.

In the Philippines it was estimated that 9.43 million dwelling units were needed between 1960 and 1980. Of these, 5.79 million were to accommodate the trends in population and average household size. As for Sri Lanka, the urban housing deficit was estimated at 120,000 units in 1962 alone. At a rate of construction around 1,200 units per year, this deficit had by 1972 increased by approximately another 75,000 units to a total of 195,000 units. In Pakistan, the urban housing shortage stood at about 600,000 units at the time of its second five-year plan (1960–65). Another expected need of 500,000 units was projected to accommodate the population increase during the period of the plan. During the third plan (1965–70), there was an accumulated back-log of 950,000 units besides the other projected need of 600,000 units to cover population increase during 1965–70.

Generally, similar conditions can be expected in countries where *per capita* gross domestic product is below U.S. $100. These will include countries such as Pakistan, Bangladesh, Afghanistan, Burma, Congo, Ethiopia, Indonesia, Korea, Laos, Nepal, Nigeria, Republic of Somalia, Sudan, Tanzania, Uganda and Yemen.

Table 84 shows the urbanization trend in recent years for Asian countries. There has already been an extensive development of shanty-towns in many large urban areas. In India, a typical urban family lives in a temporary shelter or in one room of a tenement. About 33 per cent of the households have less than 3.72 m^2 of living space per person. In Calcutta, 77 per cent of the families have less than 4.1 m^2 of living space per person.

Table 84
Rate of growth
of urban population

Country	Percentage of urban population by year				Percentage of growth within the period (years)	
China	1950E[1]	11.2	1956E	14.2	6	3.2
India	1951	17.3	1961	18.0	10	0.7
			1971E	19.9	10	1.9
Indonesia	1961	14.9	1964	11.6	3	(—)3.3
Iran	1950E	20.0	1960E	34.3	10	14.3
			1970E	40.6	10	6.3
Iraq	1950E	36.6	1960	43.6	10	7.0
			1970	57.8	10	14.2
Israel	1951E	71.7	1960E	76.7	9	5.0
			1969E	82.2	9	5.5
Jordan	1952	36.3	1961	43.9	9	7.6
Japan	1950	37.5	1960	63.5	10	26.0
			1965	68.1	5	4.6
Malaysia (West)	1957	42.7	—	—	—	—
Nepal	1954	2.8	1961	3.6	7	0.8
Pakistan	1951	10.4	1961	13.1	10	2.7
			1970E	13.1	9	0.0
Philippines	1956E	35.3	1960	29.9	4	5.4
Thailand	1956E	8.7	1960	18.2	4	9.5

1. E=estimated.

Source: U.N. Statistical Year Book, 1969.

Housing programmes

In New Delhi, town planners have introduced U-shaped rows of houses in all the sectors of Ramkrishnapuram (an area of newly built accommodations for civil servants), which provide privacy and open space. Scarcity of living space in Tokyo has impelled the Japanese Government to depart from constructing traditional single-family wooden houses, substituting four-storey flats of reinforced concrete in the outskirts of the city.

Between 1941 and 1946, 1,600 new apartments were built in Madras, India, for 458,000 new migrants. After independence, during India's first three five-year plans, about 400,000 dwelling units were constructed under various housing schemes. During 1966–69, 74,776 houses were constructed under the social housing scheme, as well as 32,802 units in Delhi under the Jhuggi and Jhonpri scheme. During 1969–72, 66,430 houses were constructed under various social schemes.

Many projects have been initiated to help solve growing needs for shelter such as in India the integrated subsidized housing scheme for industrial workers and economically weaker sections, low income group housing scheme, slum clearance improvement scheme, Jhuggi and Jhonpri removal scheme, environmental improvement of slums, village housing scheme, middle income group housing scheme, scheme for state government employees. In addition, master plans for the development of larger cities in India have been launched.

Industrialization

Successful adoption of technological innovations leading to industrialization increases total and *per capita* productivity, consumption of energy and consumption of natural resources. It also increases the rate of flow of materials throughout the economy. Perhaps the most important fact is that industrialization influences population distribution patterns and human economic and social conditions.

Social and economic effects of industrialization

As new industries develop, cities grow. Within the past 100 years, industrialization has promoted intensive urbanization of most of the advanced nations. This pattern was made possible by the increase in agricultural productivity, which in turn was facilitated by the industrialization of agriculture (mechanization, irrigation, fertilization and chemical control of pests). For example, in 1850, 75 per cent of the population in western Europe could provide surplus food for the remaining 25 per cent. This proportion of farmers and non-farmers is now reversed in some advanced areas. In France, the United States and New Zealand, countries which export food, 50 per cent of the population live in towns. Urban population constitutes 80 per cent in the United Kingdom which imports food in exchange for manufactured goods.

As it proves more and more difficult to meet the needs of urbanized and industrialized areas, problems of sewage, air and water pollution, and solid waste disposal become more and more intense. Some industrial nations, while developing economically, are faced with immense problems of environmental degradation.

Industrialization in developing countries

The growth of industries in the developing countries has a great impact on their social traditions. In India it is especially affecting the joint family system and the institution of caste. During the entire pre-industrial period, these twin institutions withstood the onslaught of political upheavals and provided such benefits as competition-free traditional livelihood, occupational efficiency, sound economy, emotional security and social welfare, and stability. These benefits of course, were accompanied by certain problems and ills. The attractions and prospects of urban life provided by industrial growth not only drew the landless and jobless people but also the energetic and enterprising well-to-do young men who were no longer content with traditional occupations. Their caste distinction and social barriers were submerged in the anonymity of the city. The growing extension of modern transportation and communication also contributed to the removal of barriers. In the rural economy individual earning had been submerged in the common family pool. In the new situation the urban-employed man remained separated from the rural home and became engrossed in his own welfare and that of his wife and children. He thought little about the joint family. Hence, this process led to the breakdown of the joint family.

Status of industrialization

Table 85 presents major technological innovations, human type and energy consumption as estimated throughout a million years of man's history. Energy consumption here includes food consumed by man and draft animals. It is apparent that increased technological advancement up to the present time has been accompanied by drastically increased *per capita* energy consumption. (See Chapter 4 on 'energy'.)

Table 85. Major technological innovations, human progress and energy consumption

Major technological innovation	Human type	*Per capita* energy consumption/day (kc)	Example	
			Place	Period
Primitive stone implements	Primitive man	2 000	Eastern Africa	1,000,000 years ago
Domestication of fire	Hunting man	4 000	Europe	100,000 years ago
Agriculture & domestication of animals	Primitive agricultural man	12 000	Fertile Crescent	B.C. 5000
Use of water, power and wind power	Advanced agricultural man	24 000	North-western Europe	A.D. 1400
Steam engine	Industrial man	70 000	England	A.D. 1875
Electronic computer	Technological man	230 000	United States	A.D. 1970

Source: Earl Cook. 'The Flow of Energy in an Industrial Society'. *Scientific American* (New York), September 1971, p. 135–6.

Gross domestic product (GDP) is another indicator of the degree of industrialization and technological advancement. However, some countries rich in mineral resources have high GDPs through extractive installations which are sometimes foreign-owned. The country may thus have a high GDP without being technologically advanced.

R. B. Sutcliffe, in his book *Industry and Underdevelopment* states that a country becomes industrialized when the industrial sector absorbs more than 10 per cent of the total population, contributes at least 25 per cent of GDP, and a minimum 60 per cent of the industrial output comes from manufacturing.

On the basis of the above criteria, Argentina, Uruguay, Japan and Israel are the only countries outside the recognized industrial regions of the world (North America, Europe, the U.S.S.R., Oceania) which are themselves industrialized. The western European countries of Cyprus, Greece and Turkey are not yet fully industrialized, though Ireland, Portugal and Yugoslavia as well as Chile are approaching full industrialization. Table 86 gives detailed economic, agricultural and industrial data for various Asian countries.

Table 86. Population, total and *per capita* GDP, *per capita* energy consumption and the share of agriculture and industry in the economy of Asian countries

Country	Population (millions) 1965	GDP (million U.S. $) 1965	GDP *per capita* (U.S. $) 1965	*Per capita* energy consumption (coal equivalent in kg)	Share of agriculture (in percentage)		Share of industry (in percentage)			
					Population (1965)	GDP	Industrial output		Manufacture output	
							Population	GDP	Population	GDP
Afghanistan	15.5	1 040	69	23	84	50 (1968)	—	—	—	—
Burma	24.7	1 575	64	48	62	34 (1967)	—	20	—	16
Hong Kong	4.0	1 097	298	625	—	—	18	—	15	—
India	473.0	46 128	95	171	60	52 (1967)	5	21	4	15
Indonesia	105.0	8 479	83	113	67	49 (1967)	3	13	2	11
Iran	23.4	6 275	256	383	49	21 (1968)	7	35	5	29
Iraq	7.2	2 302	281	504	—	—	4	47	3	9
Israel	2.6	3 181	1 241	2 237	—	—	11	35	8	25
Japan	98.0	82 090	838	1 782	24	12 (1967)	16	38	12	28
Khmer Republic	6.3	749	122	45	75	41 (1966)	2	18	1	9
Kuwait	0.5	2 095	4 411	13 943	—	—	12	—	4	—
Malaysia	8.1	2 236	278	338	55	28 (1966)	4	26	2	10
Nepal	10.1	759	75	8	92	66 (1966)	—	—	—	—
Pakistan	102.9	10 433	101	88	68	47 (1967)	4	16	3	11
Philippines	32.3	7 776	240	205	58	33 (1967)	5	26	4	19
Singapore	2.9	988	530	639	9	—	7	—	5	—
Sri Lanka	11.2	1 541	138	108	50	40 (1967)	4	13	3	6
Taiwan	12.4	2 480	199	653	47	24 (1967)	5	27	3	19
Thailand	30.6	3 559	116	128	78	31 (1967)	2	20	2	12

Sources: Unesco Statistical Year Book, 1966; *United Nations Statistical Year Book*, 1969; *Agricultural Education in Asia, a General Survey*. Paris, Unesco, 1971; R. B. Sutcliffe, *Industry and Development*. Manila, Addison-Wesley Publishing Co., 1971.

Kuwait and Hong Kong are the only Asian countries (excluding industrialized Japan and Israel) where more than 10 per cent of the population is absorbed in the industrial sector. Iran, Iraq and Malaysia, in addition to Japan and Israel, get more than 25 per cent of their GDP from the industrial sector. More than 60 per cent of the industrial output in Burma, India, Indonesia, Iran, Israel, Japan, Pakistan, Philippines, Taiwan and Thailand is derived from manufacturing.

Pace of industrialization

Within the last two decades (1950–70), population of the world has increased 1.45 times, while industrial production has increased 2.80 times and GDP 2.70. Asian population growth is about the same as that of the world, while industrial production has increased 8.20 times. Table 87 shows average annual rates of growth in population, GDP, agriculture, and industrial activity for Asian countries.

Table 87
Average annual rates of growth during 1960–68 in selected countries (in percentage)

Country	Population (1963–68)	Gross domestic product: Total	Gross domestic product: *Per capita*	Agriculture	Industrial activity (total)
Burma	2.1	3.4	1.3	4.3	4.0
India	2.5	2.8	0.3	—0.3	5.7
Indonesia	2.4	2.2	—0.2	1.7	2.1
Iran	2.5	7.6	4.7	3.2	12.4
Iraq	3.0	6.7	3.3	6.1	7.1
Malaysia	2.9	5.8	2.6	4.2	8.5
Philippines	3.5	5.2	1.7	4.9	4.9
Thailand	3.1	7.6	4.3	4.0	12.0

Source: U.N. Statistical Year Book, 1969.

Agriculture is the main industry in Asia and the rural masses constitute the major part of the population. Until rural life is progressively modernized, it will not be possible to set up a modern industrial economy. Such an economy represents a technical complex and not a conglomeration of isolated pieces of technology as is now usually seen in Asian and other developing countries. Modern industry of the West is a product of societies having a labour scarcity and as such is unsuitable for Asian nations which have abundant unused manpower. Modern green revolution technology (improved seeds, fertilizers, pesticides, irrigation, improved breeds, etc.) with a minimum of mechanization and a maximum of manpower utilization can be applied to increase the quantities and qualities of agricultural and livestock products. In addition the expansion of small industries, development of practical knowledge, and skill and work experience rather than formal education are needed. Furthermore, health facilities, water supply improvements, and better housing will contribute a great deal in progressively modernizing rural life, raising the economy of the rural masses, and accelerating the process of industrialization.

Further reading

BOURGEOIS-PICHAT, Jean. *La démographie*. Paris, Gallimard, 1971, 188 p.
ECAFE. *Interrelation between population and manpower problems*. Bangkok, Asian population studies, No. 7, 1971.
——. *The regional seminar on population aspects of social development* (1972). New York, Asian population studies, No. 11, 1972.
FREJKA, T. *The future of population growth*. The Population Council, 1974, 280 p.
FRYER, D. W. *Emerging South-East Asia: a study in growth and stagnation*. New York, McGraw-Hill, 1970.
HEER, David M. *Society and population*. Englewood Cliffs, New Jersey, Prentice Hall, 1968.
KEYFITZ, Nathan; FLIEGER, Wilhelm. *Population: facts and methods of demography*. San Francisco, Freeman, 1971.
MOUNTJOY, Alan B. *Industrialization and underdeveloped countries*. University Library of Geography, 1966.
MUKERJEE, R. *Man and his habitation*. Bombay, Bhatkal for Popular Prakashan, 1968.
MYRDAL, Gunnar. *The role of science and technology in the development of underdeveloped countries in South Asia*. Bombay, Asia Publishing House, 1964.
PETERSEN, William. *Population*. New York, Macmillan, 1969.
POPULATION COUNCIL. *Country profiles*. A review without regular periodicity, since July 1970 (South Korea, Hong Kong, Indonesia, Iran, Israel, Japan, Malaysia, Mauritius, Nepal, Pakistan, Philippines, Taiwan, Thailand, Turkey). New York, The Population Council.
SCOTT, Greer *et al.* (ed.), *The new urbanization*. New York, St. Martin Press, 1968.
SOUTH-EAST ASIA DEVELOPMENT ADVISORY GROUP OF THE ASIA SOCIETY (SEADAG). *Reports*. Washington, 1972.
——. Population growth and development in South-East Asia (Taeuber, J. B.).
——. Population growth in South-East Asia: a summary review (Whitney, V.).
——. Implication of prospective urbanization for development planning in South-East Asia (Jones, G. W.).
——. Rural fertility growth in West Malaysia (Palmore, J. A.).
——. Trends and implications of urban growth in Malaysia (Ishak bin Ariffin).
——. Population growth and its implication in Indonesia (Iskandar, N.).
——. Some consequences of population growth in Java (Singarimbur, M.).
SOUTH-EAST ASIA DEVELOPMENT ADVISORY GROUP OF THE ASIA SOCIETY (SEADAG). Socio-environmental aspects of rapid urbanization in Indonesia, with special reference to Jakarta (Suharso).
——. Fertility, social class and outmigration from two rural communities in the Philippines (Hendershot, G. N.).
——. Coping with internal migration in the Philippines: problems and solutions (Laquian, A.).
——. Analysis of urbanization in Japan (Inoue, S.).
——. Dualism in the Asian city: the implications for city and regional planning (McGee, T. G.).
——. Trends and implications of urban growth in Thailand (Nibondh Debavalya).
——. Research on demographic, social and economic change in Thailand (Visid Prachuabmoh).
THOMLINSON, Ralph. *Population dynamics: cause and consequences of world demographic change*. New York, Random House, 1965.

Part IV

Reproduction in human populations

Preamble

The human life span is about 120 years. The oyster has a life span of 12 years, while a toad lives 35 years, a mouse 39 months, and a bristle cone pine tree 4,600 years. Life spans are the confirmed maximum survival times characteristic of a species. Technically only one individual in the history of the species will have survived the record time described as the life span. All the others die somewhere between the instant of live birth, and death at the oldest age known.

For every population of a species there is an average length of life which individuals within that population can expect. This life expectancy varies among populations. Wild birds in the field live a considerably shorter time than populations of the same species of bird kept in captivity. At the present time new born females of the Netherlands have a life expectancy of 76 years, while members of other present human populations, living without the benefit of modern technology and medicine, often have life expectancies at birth that do not extend beyond 30 or 40 years.

Like all species of plants and animals, the human works hard at staying alive and the human body struggles against its inevitable death. Every individual dies when it can no longer carry out any one of the life processes which maintain it. Much can be done to extend the life of individuals through programmes of health and safety. But ultimately the organism succumbs to the enfeebling effects of old age. The phenomenon of aging is common to all species and no individual, regardless of its constitution, can escape the fact that it can live only so long. If life on earth depended upon the ability of each organism to remain alive indefinitely, life would soon disappear. In 120 years, the human would certainly become extinct—as would elephants in 77 years, rabbits in 14 years, and wheat in one year.

The life process of reproduction enables each species to exist beyond the life span of its healthiest individual. Reproduction provides continuity and life to the species just as other life processes support the lives of individual organisms. The loss of the ability to reproduce will not significantly affect the individual's ability to sustain itself, but should that loss occur among a critical portion of a population of a species (caused by disease, loss of breeding grounds, change in climate), that population will

die out, become extinct. Thus, reproduction is the process by which species are maintained.

Among very many species of simple plants and animals, single individuals give rise to new, separate and complete organisms, without the participation of any other individual. This one-parent reproduction is essentially a growth process in which somatic cells separate from the body of the parent and develop into offspring genetically identical to the parent. A sexual reproduction takes such forms as fission, budding, sporulation and fragmentation. As a reproductive process it is well suited to simple organisms which live in stable environments where variability is not essential to survival. All environments do change, however, and reproduction must also be accomplished in a way that assures variability. This variability is reached when genetic exchanges take place during differentiation of sex cells. Therefore, sexual reproduction involving the union of two gametes results in individuals genetically different from paternal phenotypes. Sexual reproduction is a complex process requiring the development in both sexes of structures which produce and transport cells that will come together and unite in a liquid medium at the right place and the right time. For terrestrial organisms the sexual processes become even more complicated with the continuing need to maintain a liquid environment for fertilization and embryonic development.

In the human, sexual reproductive complexity is further compounded by culture. We humans create profound changes in our own environments. Religion, economics, politics, education, technology impose demands and restraints on our reproductive behaviour. The reproductive relationships between the sexes, parents and offspring, the individual and society (population) are dramatically modified by what our culturally created environment directs us to do. Thus, for the human, it is highly appropriate to look at reproduction both biologically and sociologically.

The objectives in this section are to examine, first, some biological facts of human reproduction, and, second, some of the effects of human culture on reproductive behaviour.

16 Human reproduction

Differentiation of the male and female reproductive systems from indifferent origins

Even though the sex of an individual is genetically determined at the time of fertilization, it is not possible to distinguish males from females in early stages of development. Inspections of the internal and external genitalia prior to the eighth week of human development cannot reveal the ultimate sex of the individual. Sex can be determined by examination of the chromosomes.

The genital and urinary systems begin to form in close association during the fifth week of development. Germ cells, originating in yolk sac endoderm, migrate by ameboid movements to a pair of genital ridges which bulge into the coelom. Closely associated with each genital ridge at this time is a functioning primitive kidney, the mesonephros, complete with tubules and ducts. During the sixth week of embryological development, a second pair of ducts, the Mullerian ducts, form from a groove on the (uro) genital ridge. By the seventh week the tissues of the genital ridge furrow, and separate from the primitive kidney, forming a pair of gonads of undisclosed sex. These gonads could become either testes or ovaries. This is the stage of the indifferent reproductive system during which the embryo possesses a pair of undifferentiated gonads and two sets of ducts. If the genetic blueprints direct that the individual is to be a female the gonads become ovaries; the Mullerian ducts become the oviducts, uterus and vagina; and tubules and ducts of the mesonephros become rudimentary structures. If the embryo is male, the gonad becomes the testes, the mesonephric tubules and ducts are appropriated for functional use as ducts for the discharge of sperms, while the Mullerian ducts become rudimentary. Thus when sex is differentiated, the duct system appropriate to the gonad develops while the complementary set of ducts regresses. Fig. 114 illustrates the transformation of the internal indifferent reproductive system to normal male and female systems.

While the internal genitalia are differentiating (more rapidly for males) similar developments are occurring with the formation of external genitalia. Late in the fifth week a genital tubercle appears between the tail and the umbilical cord. The underside of the genital tubercle becomes grooved between two uro-genital folds. As the tubercle enlarges forming a phallus, its tip forms the rounded glans, and a foreskin

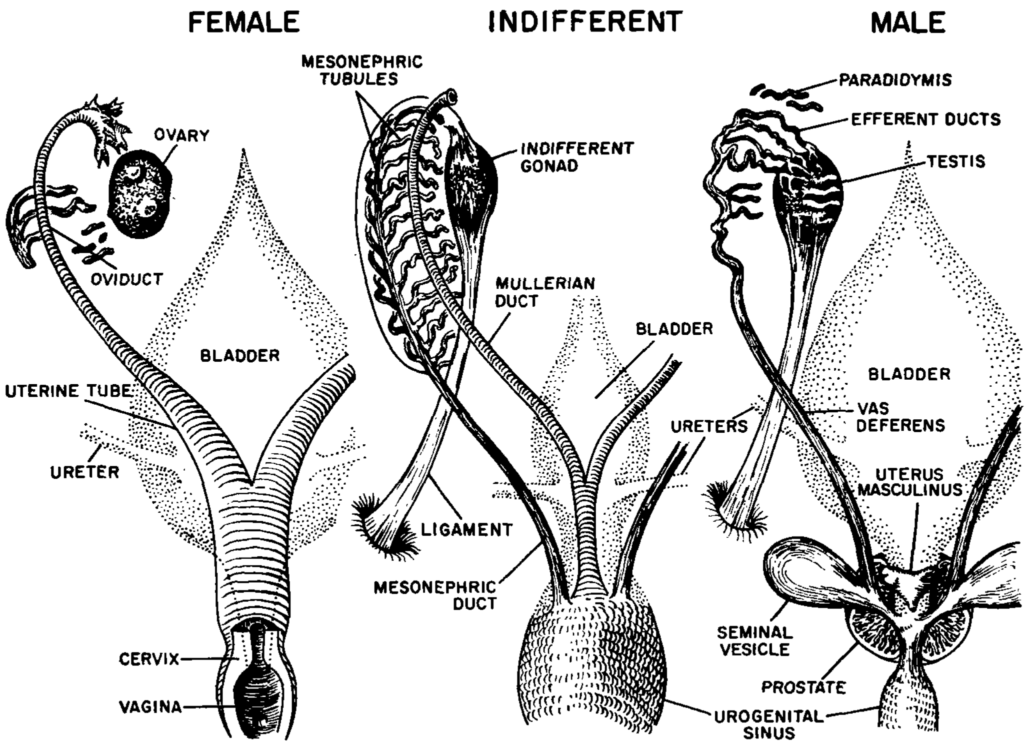


Fig. 114
The internal indifferent reproductive organs and their development into male and female reproductive organs in the embryo
Source: Nalbandov (1964). *Reproductive Physiology*. Freeman, 1964.

(prepuce) that covers the glan forms from the urogenital folds. Subsequently, the tissues lateral to the enlarging genital tubercle enlarge forming the genital swellings. If the embryo develops into a male, the genital tubercle enlarges to become the penis and the genital swellings form the scrotum. In the developing female the genital tubercle becomes the clitoris, the genital folds become the labia minora, and the genital swellings turn into labia majora. Details of the differentiation of the external genitalia can be seen in Fig. 115.

The reproductive structures of males and females develop from structures that are common to individuals of both sexes in the embryonic stages. Aside from one's fascination with the evolutionary significance of the differentiation of the reproductive system, knowledge of the phenomenon is useful in understanding common variations of reproductive anatomy and physiology. Table 88 summarizes the homologies of the male and female reproductive systems, and notes the functional and rudimentary

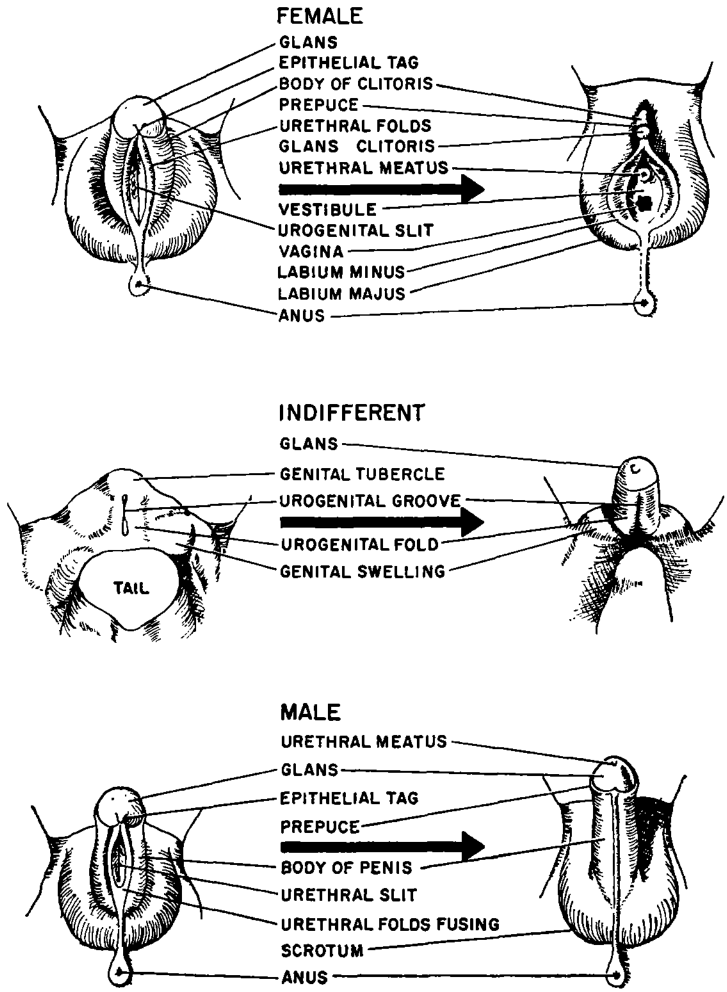


Fig. 115
The external indifferent genitalia and their development into male and female genitalia

structures of the indifferent sex, the male, and the female. Not uncommonly the rudimentary structures of the opposite sex may differentiate abnormally under the effects of hormones, may lead to disorders of the reproductive system.

Experimental studies with rabbits and mice from which indifferent gonads were removed suggest the mechanism by which sexual differentiation occurs. After the indifferent gonads were removed the internal and external genitalia became structurally feminine. This occurred whether the individual was genotypically male or female. The removal of testes from the bodies of embryos prior to the completion of the differentiation of the male genital structures leads to the formation of feminine structure

Table 88
Homologies of male and female reproductive systems

Indifferent	Male	Female
	Internal genitalia	
Gonad	Testis	Ovary
	Rete testis	Rete ovarii[1]
Mesonephric tubules	Vas efferens	Rudimentary structure
Mesonephric duct	Epididymis Vas deferens Ejaculatory duct Seminal vesicle	Rudimentary structure
Mullerian duct	Appendage of testis[1]	Fimbria of oviduct Oviduct Uterus Vagina (all or part?)
Urogenital sinus	Prostatic, membranous, and cavernous urethra	Urethra Vestibule Vagina (in part?)
	Bulbo-urethral (Cowper's glands)	Vestibular glands (Bartholin's glands)
	External genitalia	
Genital tubercle	Glans penis	Glans clitoris
	Corpus penis	Corpus clitoridis
Urethral folds	Raphe of scrotum and penis	Labia minora
Labiscrotal swellings	Scrotum	Labia majora

1. Rudimentary.

Source: Nalbandov. *Reproductive Physiology*, Freeman, 1964.

in parts not yet differentiated. It thus appears that embryonic testes impose masculinity and repress femininity on the bodies of rabbits and mice, while the ovary does nothing to effect the differentiation of the female genital tract.

The male reproductive system

Although formed at the same site as the ovaries, the testes descend into a partitioned skin covered sac, the scrotum, during the sixth or seventh month of development or shortly after birth. The testes are suspended in the scrotum by a spermatic cord which encloses sperm ducts, blood vessels and nerves. Occasionally, the testes fail to descend into the scrotum, a condition known as cryptorchism. At body temperature the testes fail to produce living sperm and individuals with undescended testes are sterile. Normally, the scrotum is sensitive to heat and will relax when warm and contract when cold, raising the testes against the warm abdominal wall.

Structurally a testis is a mass of coiled (seminiferous) tubules which are arranged

in partitioned sections. At puberty the seminiferous tubules become hollow, and begin to produce great numbers of the sperm. The sperm pass into the rete testes before being collected by ducts external to the testes themselves. Between the seminiferous tubules are the interstitial cells of Leydig which produce, and secrete into the blood the male hormone, testosterone. This hormone is responsible for the stimulation and maintenance of masculine characteristics. One of the functions of testosterone is the stimulation of protein anabolism which promotes the growth of skeletal muscle and bone.

On the back and top of each testis is the coiled epididymis which, if stretched out, is twenty feet in length. The function of this duct is to collect and store sperm, and also to secrete some of the seminal fluid in which sperm are suspended. The tail of the epididymis is continuous with the vas deferens. This duct passes upward over the brim of the pelvis and into the pelvic cavity. Small canals connect each of the two vas deferens to each of two seminal vesicles located between the urinary bladder and the rectum. The seminal vesicles are pouches from which the sperm derive nutrients and gain mobility. Contrary to popular belief, it is doubtful that the seminal vesicles store sperm. Beyond the confluence of the seminal vesicles and the vas deferens are two ejaculatory ducts which empty into the urethra. Except during sexual excitement the ejaculatory ducts remain closed, preventing the passage of sperm and seminal fluid into the urethra. The prostaste gland completely encircles the ejaculatory ducts and the urethra. Its function is to secrete into the urethra a thin alkaline additive to the seminal fluid. This alkaline constituent of the seminal fluid tends to neutralize the acidity of residual urine in the urethra of the males and the passages of the female, and thus assure the vitality and mobility of the sperm. The prostate gland enlarges with the age of the man and in later years may squeeze the urethra and cause difficulties with urination. Below the prostate gland are the two pea-like bulbo-urethral glands which secrete an alkaline lubricant into the urethra during sexual excitement. The urethra is the common duct for urine and semen (sperm suspended in seminal fluid).

The structure which contains the urethra and delivers the semen into the reproductive passages of the female is the penis. It is composed of three cylindrical masses of cavernous tissues enclosed in separate fibrous coverings and bound together by skin. The two larger upper cylinders are the corpora cavernosa penis. The smaller lower cylinder, the corpus cavernosus urethra or corpus spongiosum, contains the urethra and terminates as the slightly bulging and sensitive glans. The three cylinders form a triangle. A retractable foreskin, the prepuce, covers the glans and is sometimes surgically removed by circumcision. Two muscles are associated with the penis; one causes the expulsion of urine and seminal fluid, the other prevents blood from being drained through the veins while the penis is erect. During sexual excitement, arteries of the penis dilate and because drainage is hindered the spaces in the spongelike tissue engorge with blood causing the penis to become enlarged, hard and erect. The shape, which is cylindrical in the flacid state, becomes more angular, approaching the form of a triangular prism. Over-all the penis will increase 50 per cent in length, 40 per cent in circumference and 30 per cent in diameter. Erection enables the insertion of the penis into the vagina of the female during sexual intercourse.

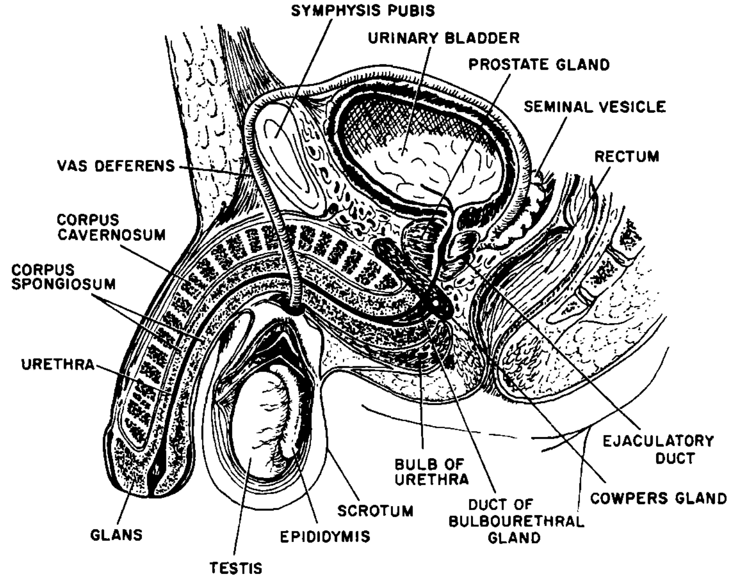


Fig. 116
Sagittal section of the human male pelvis showing reproductive organs

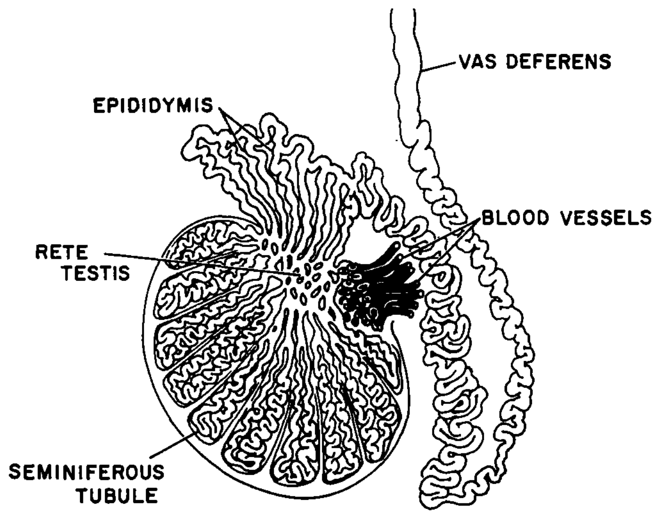


Fig. 117
Section through testis showing arrangement of ducts

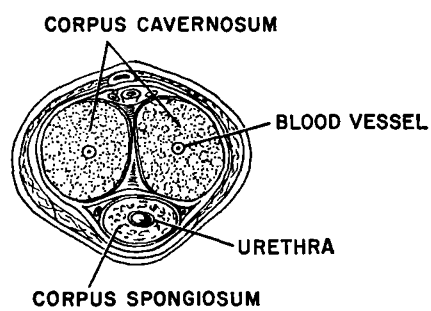


Fig. 118
Cross section of penis

Figs. 116, 117 and 118 provide various views of the structures of the male reproductive system.

Fertility of the male depends upon factors including his ability to achieve erection (potency), provision of proper nutrients and maintenance of neutral to alkaline pH in the seminal fluid, a volume from 3 to 6 cc of semen containing 60 million sperm per cc with 60 per cent of the sperm mobile two hours after ejaculation, and with 60 per cent of the sperm showing normal morphology. Low sperm counts sometimes can be traced to present or past diseases (e.g. mumps, venereal disease), excessive coitus, radiation exposures, heat, tensions or heavy metal poisoning. In treating male infertility, the physician studies the life history of the man, and, when indicated, prescribes changes in life style. For example, since too frequent intercourse reduces sperm counts, and long periods of continence, though increasing sperm counts, will reduce the degree of sperm mobility, the physician may recommend regular coitus at a frequency of two to three times per week; or where the man's testes are exposed to high and prolonged heat as in wearers of tight underwear, which tends to suppress spermatogenesis, the physician will recommend appropriate changes. When male infertility appears to be caused by an endocrine disorder, hormonal and vitamin treatments thus far tried out have not proven to be very effective, and much continued research in this field is necessary. Approximately 10 to 15 per cent of married couples have barren relations, and within this population about 30 to 35 per cent of the barrenness is attributed to the male partners.

The female reproductive system

The ovaries, producing eggs and hormones, are the female counterparts of the testes. They lie deep in the pelvic cavity one on either side of the uterus, and are each suspended from above by a ligament. They are no larger than the small toe and weigh between two and three grammes. Nevertheless, an ovary of a new-born girl may contain as many as one-half million undeveloped eggs. That number diminishes with age, and after menopause all eggs disappear. The immature eggs, or oocytes, are surrounded by a single layer of cells forming bag-like structures termed the primary follicles. These develop into secondary follicles containing additional cell layers around the immature egg.

Beginning at puberty one oocyte ripens every month, the ovaries alternating in irregular and unpredictable sequence. The secondary follicle, under the influence of hormones enlarges from about 100 μ to about 1 centimetre and migrates around the ovary. The follicle enlarges and forms a space (the antrium) between the egg and the follicular wall, which is filled with a fluid. The fully developed follicle is called a Graafian follicle. When fully developed the ovum (egg) encased in cells bursts through the follicle, somewhat like a rupturing boil, and is expelled into the abdominal cavity very close to one of the Fallopian tubes. At this stage the ovum is about 200 μ in diameter. The follicle then becomes the yellow corpus luteum. Should an ovum be fertilized and become embedded in the wall of the uterus, the corpus luteum will persist

throughout the pregnancy. If pregnancy does not occur, the corpus luteum, after fourteen days, becomes a white scar known as the corpus albicans.

At ovulation the released ovum enters the funnelled opening of one of the two Fallopian tubes or oviducts, which at this time may be embracing the ovary. The ovum then proceeds down the ten-centimetre-long oviduct by peristaltic action and by a current set up by ciliated cells in the lining. The unfertilized ovum lives probably for less than one day and begins to degenerate before passing completely through the oviduct.

The oviducts are attached somewhat like horns to the body of the pear-shaped uterus, which is tilted forward and held in place by eight ligaments. In gross structure, the uterus consists of a body which tapers to a lower section, the cervix. The layers of tissue of the uterus reflects its function: a vascular inner layer, the endometrium of spongy tissue that periodically enlarges and is sloughed, and during pregnancy provides a nutritive nest for the embryo; a thick muscular layer that provides strength and the force required to expel a child at birth; and a tough coat of connective tissue which binds the expansible organ together. The uterus meets the vagina at an angle of 90° above and behind the urinary bladder and communicates with the cavity of the vagina at the cervix.

The vagina lies between the bladder, urethra and the rectum. Its very elastic wall internally lined with a mucous membrane usually remains collapsed. Below the cervix it forms a circular recess. The vagina functions to discharge the menstrual fluid and to accommodate and lubricate the penis during intercourse. During parturition it serves as the birth canal.

The external genitalia of the female are collectively termed the vulva. The mons pubis is the fatty mound of tissue that overlies the pubic bone and which, from puberty onward, is covered with hair. The labia majora are the two 'large lips' which cover the

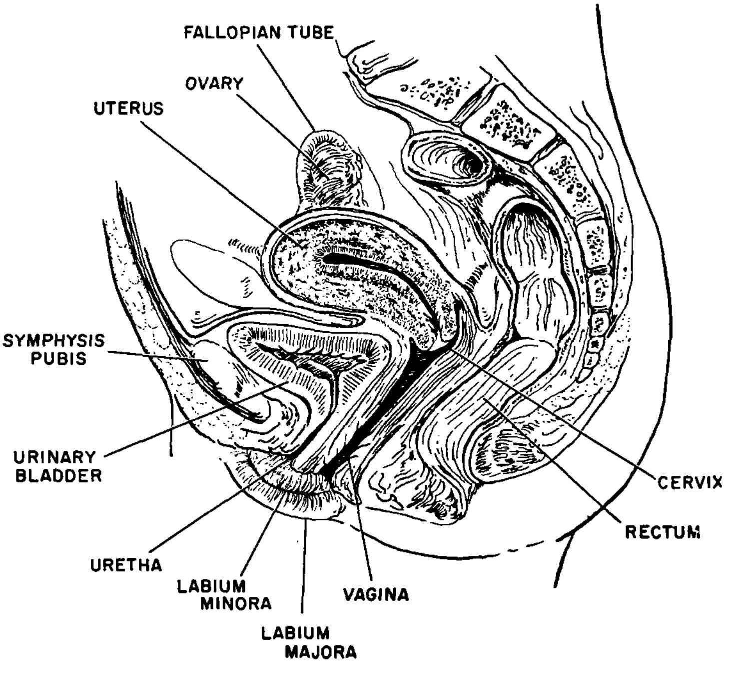


Fig. 119
Sagittal section of the human female pelvis showing reproductive organs

underlying structures while the female is standing. Beneath these a smaller pair of lips, the labia minora, enclose the opening of the vagina and the urethra. The space between the labia majora and the labia minora is the vestibule. The clitoris, a small organ homologous with the male penis, is located at the front apex of the labia minora. Behind the clitoris and enclosed by the labia minora are the openings of the urethra and the vagina. In the virginal state these openings are about the same size. A thin membrane, the hymen, normally partially covers the vagina. During the first intercourse the hymen is torn and after this it tends to regress and disappear. (Sometimes the hymen is surgically removed to avoid tearing and bleeding.)

After sexual intercourse the opening to the vagina becomes noticeably larger. On either side of the vaginal opening and at the base of the labia majora are the glands of Bartholin which secrete a mucous fluid during intercourse. However, most of the fluids produced during sexual excitement, once attributed to these glands, are now known to come from the vaginal lining.

Fig. 119, 120, 121 and 122 illustrate the major structures of the female reproductive system.

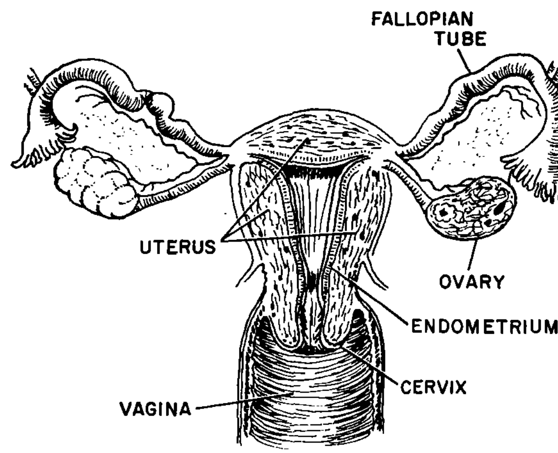


Fig. 120
Face view of internal female reproductive organs

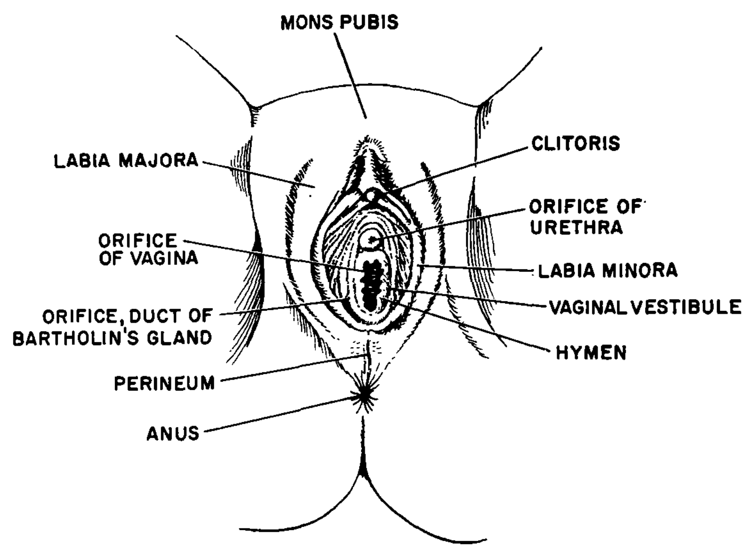


Fig. 121
External genitalia of mature female

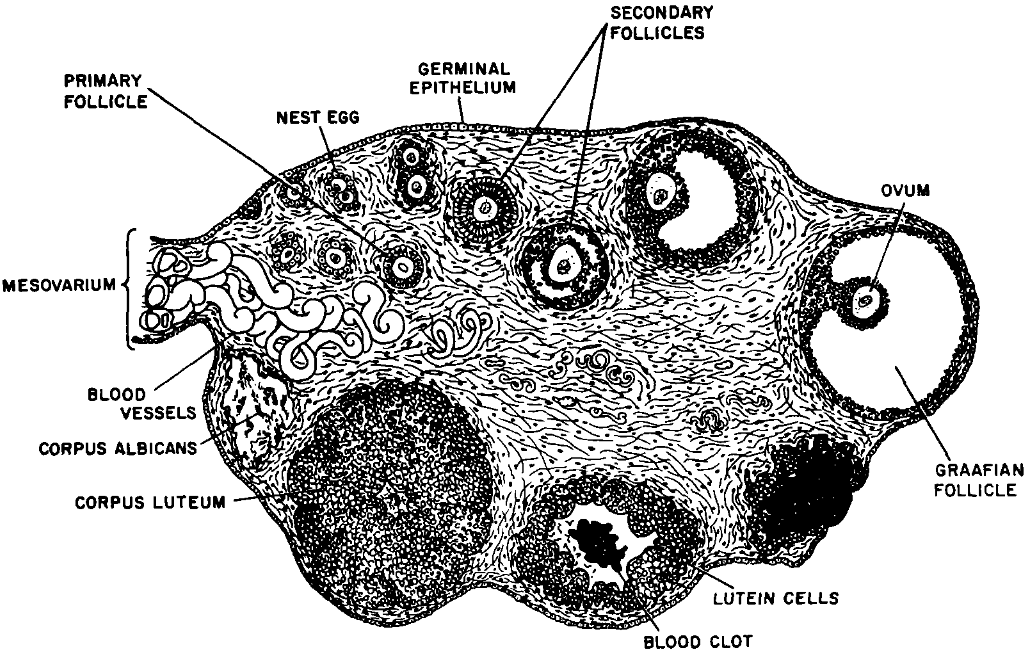


Fig. 122
Section through human ovary

Hormonal control of reproduction

Reproductive processes are controlled by a number of hormones that circulate in the blood in very small amounts, often making their identification and quantification difficult. The hormones which exert a major role on reproduction are of two general types, polypeptide hormones and steroid hormones. Most of these polypeptide hormones are synthesized and secreted from the anterior lobe of the pituitary gland and named gonadotrophins because of their main action on gonadal tissues. The sex steroid hormones are produced in ovaries and testes. Sex steroids may also be produced by the adrenal cortex and placental tissue of pregnant females. However, the amount of sex steroids, being secreted from organs of extra-gonadal origin are relatively very low.

The gonadotrophic hormones are:

a. *Follicle stimulating hormone* (FSH)

This hormone is produced in the anterior pituitary. In women it acts to stimulate growth of the ovarian follicle and, together with LH, to cause ovulation. In man FSH acts with the other pituitary gonadotrophin, LH, to stimulate testicular development and sperm formation.

b. *Luteinizing hormone* (LH)

LH is sometimes called interstitial cell stimulating hormone (ICSH). It is produced by the anterior pituitary. In women it acts with FSH to cause ovulation. This hormone

also acts with FSH in males to stimulate interstitial cell development and testosterone production.

c. *Human chorionic gonadotrophin* (HCG)

HCG is the only known normal human gonadotrophin which is not produced from the anterior pituitary. It is produced by the trophoblastic tissue of the placenta of pregnant females. Gonadotrophins of extrapituitary origin are not common among mammals. Only primates and animals in the horse family are known to produce extrapituitary gonadotrophins. HCG exerts an action very similar to that of LH of pituitary origin.

Sex steroids which exert their major roles on reproduction are:

a. *Estrogens*

They are a group of steroids which have eighteen carbon atoms in the molecule. (See Fig. 136.) Among these, estradiol-17B is the most active estrogen secreted by the human ovary. It acts to stimulate uterine endometrial proliferation and the female secondary sex characteristics (breast development, enlargement of the hips, pattern of distribution of pubic hairs). It also acts to control pituitary FSH and LH secretion through a complex feedback system.

b. *Progestins*

They are a group of twenty-one carbon steroids secreted mainly by the corpus luteum of the ovary. Progesterone is the major product of the human corpus luteum. This hormone acts in concert with estradiol to transform the estrogen-stimulated uterine endometrium into a secretory endometrium and to control through a feedback system pituitary FSH and LH secretion.

c. *Androgens*

They are a group of nineteen carbon steroids secreted mainly from the interstitial tissue of the testis. Testosterone, the major androgen, acts to stimulate male secondary sex characteristics (i.e., muscular strength, low voice, enlargement of the Adam's apple, appearance of beard and moustache, pattern of distribution of the pubic hairs) and, in feedback, to control pituitary LH secretion.

Sex steroids, as well as gonadotrophins function mainly after the onset of puberty. It is well established that the central nervous system, particularly the hypothalamus, controls the secretion of pituitary FSH and LH by producing neurohormones named FSH releasing factor (FSH-RF) and LH releasing factor (LH-RF). How and when these neurohormones act depends on the delicate balance between levels of sex steroids in the blood, the non-hypothalamic central nervous system, as well as such exteroceptive stimuli as light and environmental temperature. Fig. 123 and 124 show a recent concept of the interrelationships between the central nervous system, the anterior pituitary and the gonads in humans. One of the major differences between male and female hypothalamic regulation of the pituitary-gonadal function is that the female has hypothalamic centres controlling both tonic and cyclic release of gonadotrophins while the male has only the tonic centre. As a result, the male can secrete FSH and LH at nearly a constant rate throughout his entire reproductive life. In contrast to the male, gonadotrophin secretion in the female varies considerably during the various phases of the menstrual cycle.

The reproductive life of the female begins at puberty when the anterior lobe of

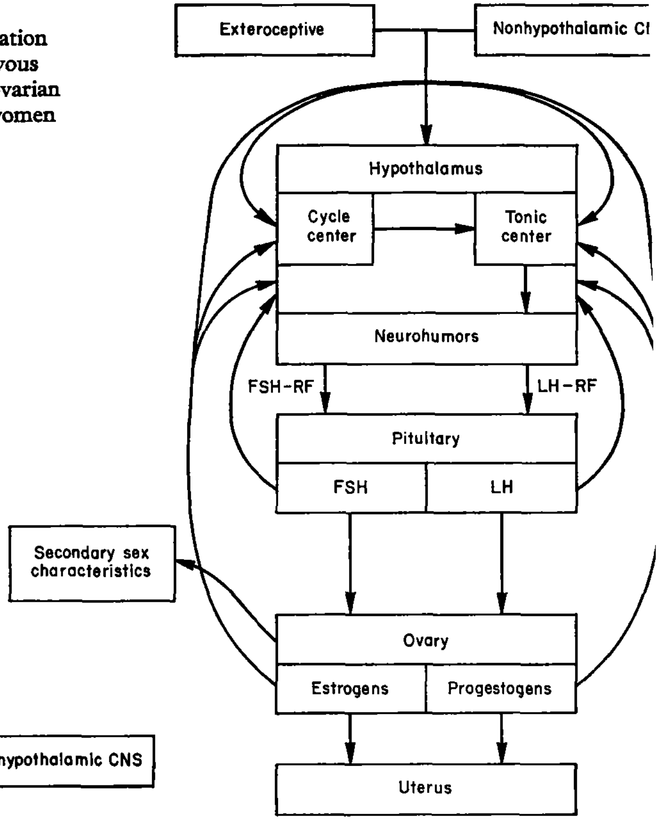


Fig. 123
Schematic presentation of the central nervous system-pituitary-ovarian interrelations in women

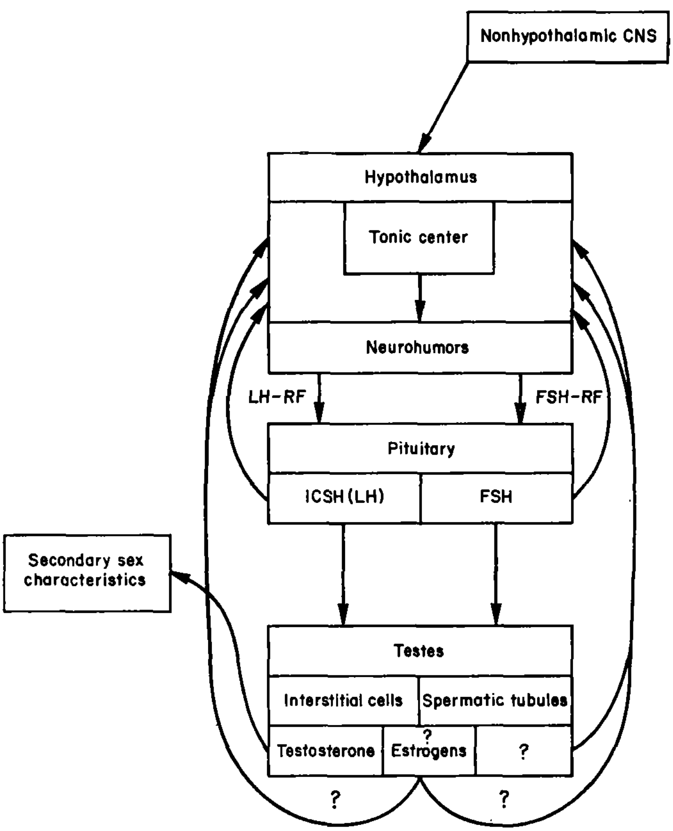


Fig. 124
Schematic presentation of the central nervous system-pituitary-testicular interrelations in men

the pituitary gland, after stimulation from the hypothalamus, initiates through chemical action such observable events as the enlargement of the breast and menstruation. Internally the successive ripenings of ovarian follicles, their transformation into glandular corpus lutea, the periodic thickening and sloughing of uterine endometrium and the fluctuating blood levels of FSH, LH, estradiol and progesterone can be observed and measured. The events are interrelated and represented in diagrams in Fig. 125.

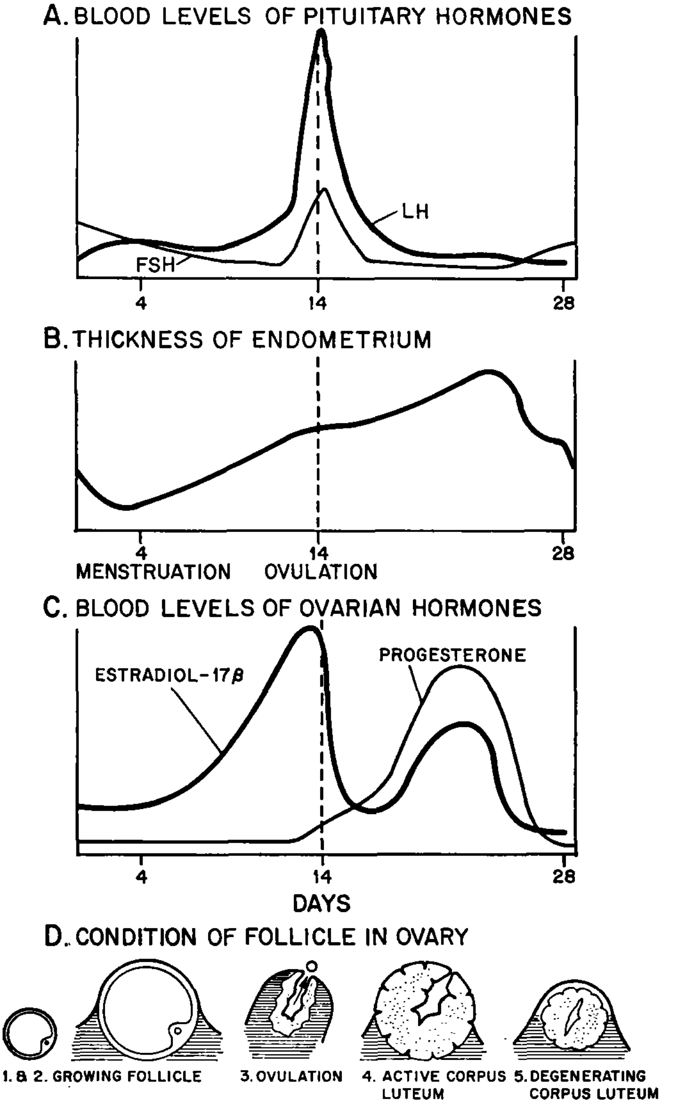


Fig. 125
Diagrams illustrating the sequence of various events in the menstrual cycle

First of all, the anterior pituitary gland secretes the hormone FSH, which presumably initiates follicular growth and development. FSH then falls to lower levels and is fairly constant until the peak at mid cycle, coincident with LH peak. This LH-FSH peak causes the discharge of the ovum from the now mature Graafian follicle. FSH then again falls to low levels remaining constant until about the 26th day of the cycle. LH is low during the early follicular phase and increases gradually throughout the follicular phase. It peaks sharply to cause ovulation at mid cycle and then falls to levels below those seen in the follicular phase and remains fairly constant. Estradiol and progesterone appear to be secreted in small amounts during the early follicular phase. However, the amount of estradiol gradually increases during the follicular phase and peaks concomitantly with the LH-FSH midcycle peaks.

Once ovulation has occurred, the yellowish corpus luteum is formed. Although blood concentrations of LH and FSH are low during the luteal phase, it has been shown that small amounts of LH are required for normal corpus luteal function. The corpus luteum secretes the hormone progesterone which reaches its highest levels during the luteal phase. It also secretes estrogen but in smaller amounts than were produced by the Graafian follicle. While the above events are occurring, the mucous membrane lining the uterine wall, the endometrium, thickens to a depth of from two to three millimetres under the influence of rising estrogen levels during the follicular phase and increases to a thickness of from four to six millimetres during luteal phase. Estrogen stimulates the regeneration of the lining which had just been stripped away; the progesterone causes a secretory change providing nourishment for the embryo prior to implantation. If the egg, now in the Fallopian tube, is not fertilized, the corpus luteum desintegrates, and the progesterone and estrogen hormone levels decline. The sudden fall of the level of progesterone and estrogen triggers a breakdown of the endometrium. Small tightly coiled arteries constrict, cutting off the blood supply to the tissues, leading to the destruction and sloughing of the inner layers of the endometrium. The expulsion of this tissue and blood through the vagina is menstruation. Conventionally, the normal 28-day menstrual cycle is counted from the onset of menstrual flow, so that by referring to Fig. 125, such events as ovulation, endometrial tissue build-up and peaking levels of estrogen, progesterone, FSH and LH can be observed against a 28-day co-ordinate.

It is not certain what actually causes the corpus luteum to shrink and stimulate FSH to initiate the cycle all over again. One explanation suggests that rising levels of progesterone inhibit LH production and subsequent drops in estrogen and progesterone levels actuate FSH production in the feedback system. Current evidence suggests that the role of the brief and heavy dose of LH is merely to effect ovulation and initiate the process that causes the corpus luteum to persist in the event fertilization occurs. The fertilized egg itself secretes the powerful hormone, chorionic gonadotrophin, which signal the corpus luteum to continue the production of progesterone and estrogen, thus ensuring the maintenance of the endometrium and preventing the production of FSH and the ripening of further eggs. Starting at the third month of pregnancy, the placenta also produces estrogens and progesterone.

Birth-control pills contain estrogen and progesterone-like substances. They enable

the blood to maintain high estrogen and/or progesterone levels preventing the development of follicles and ova through the suppression of the pituitary gland. There is evidence that the suppressive action is actually on the hypothalamus which regulates pituitary activity. The effect is to simulate pregnancy but to allow the menses to occur at the normal time by discontinuing the administration of the hormones several days prior to the menses. The action of birth control pills will be discussed further in Chapter 19 of Part V, page 413.

The chief causes of infertility in females can be associated with cervical, uterine, tubal and hormonal factors. The cells of the cervix produce a mucus containing sugars and amino acids essential to the metabolism and migration of the sperm. The secretion of cervical mucus is controlled by hormones and affected by infections. Treatment for infertility related to cervical factors may be the administration of estrogen which regulates the production of mucus and induces specific changes in cervical anatomy throughout the cycle, or medication against a genital infection. Sometimes the endometrium of the uterus will not develop adequately due to insufficient production of progesterone by the corpus luteum. In such cases of infertility progesterone or a substitute may be administered during the luteal phase of the cycle. For about 30 per cent of all infertile couples, infertility can be traced to blocked Fallopian tubes for which there are numerous treatments including surgery, insufflation, hormones and anti-infection medication. Of the more dramatic treatments for female infertility are those which attempt to intervene in the reproductive cycle with hormones with which the patient is deficient. Women whose pituitary glands fail to release sufficient FSH to stimulate the development of follicles may be treated with FSH, extracted from cadavers or from the urine of menopausal women whose ovaries do not respond to their pituitary hormones with inhibiting estrogen. After follicular growth has progressed and the egg is ripe, chorionic gonadotrophin, which acts like LH, is administered to stimulate ovulation. Current problems with this procedure are the expense and rarity of the hormones, and the uncertainties in controlling the number of eggs that ripen and shed within any one menstrual period. Multiple births are commonly the result of overdosage. The discovery of the FSH-LH releasing factor, isolated from the hypothalamus, and the chemical synthesis on an industrial scale have provided a new way to induce ovulation through the stimulation of the woman's pituitary to release its own gonadotrophins—FSH and LH. Another drug, clomiphene, apparently acts one step up the chain of hormonal action by providing the initial signal to the hypothalamus to activate an otherwise non-responsive pituitary and thus start the hormonal chain of events.

Coitus

The sex cells of humans, like most terrestrial animals, unite within the genital tract of the female. This is accomplished by the sexual embrace, coitus or copulation, during which the penis of the male is introduced into the vagina of the female followed by the ejaculation of sperm in seminal fluid.

Physiological studies of the sexual responses of males and females have revealed a pattern of four phases: excitement, plateau, orgasm and resolution. This pattern was induced from such measurements and observations during copulation as blood pressure, pulse, muscular tension, breathing, vaso-congestion, sweating and dilations. The excitement phase can be initiated by a wide variety of physical and psychological stimuli. If the stimulation is appropriate to the individual, the rate of increase of sexual tension is rapid. This phase can be greatly prolonged or aborted. The plateau phase is characterized by levels of sexual tension from which the individual may move to orgasm. If either the male or the female fail to achieve orgasm, a prolonged period for recovery occurs. The orgasm is the sudden involuntary release from sexual tension. While males appear to conform to standard patterns of ejaculatory response with little variation, the female orgasm aside from basic physiologic reactions appears to be affected by psycho-social factors. The resolution phase carries the individual back to the unstimulated state. During this phase women have the capacity to return with proper stimulation to the orgasmic phase, while the male experiences a refractory period during which he is incapable of re-excitation.

Throughout sexual intercourse, changes in the anatomy of the female have been observed. The vagina corrugates, possibly acting as a frictional aid to the penis, the vagina enlarges below the cervix forming a cavity in which the pool of seminal fluid forms, the clitoris, labia minora and uterus become erect, and after the orgasm the mouth of the uterus appears to open. The hormone oxytocin which stimulates muscle contractions in the uterus during labour has been detected during coitus. One may assume that many of the responses of structures to coitus are adaptations that tend to ensure the fertilization of egg by sperm.

Fertilization

Fertilization involves the fusion of an egg and a sperm producing a zygote which then undergoes cell division, and proceeds to develop and grow into an embryo.

The 'pay load' of a sperm cell (see Fig. 126) is the head which contains the genetic contribution of the father. Covering the head is a cap, the acrosome, which contains an enzyme enabling the sperm to penetrate the layers of cells that cover the ovum. The mid-section of the sperm is the site where sugars in the seminal fluid are metabolized, providing energy to the tail. The tail of the sperm articulates with the head at the neck where an effective ball and socket type joint functions. The movement of a sperm cell from the vagina and its capacity to fertilize successfully an egg are dependent upon such variable factors as age of sperm, age of ovum, pH of media, temperature, oxygen level, volume of semen, numbers of sperm cells, location of ovum, capacitation (perforation of the wall of acrosome and release of enzyme), viscosity of uterine mucus, and the possible presence of antibodies which immunize the female against sperm antigens.

The egg is a cell which becomes extremely large because of the accumulation of nutritive substances called vitellus or yolk. It contains, just prior to fertilization, its

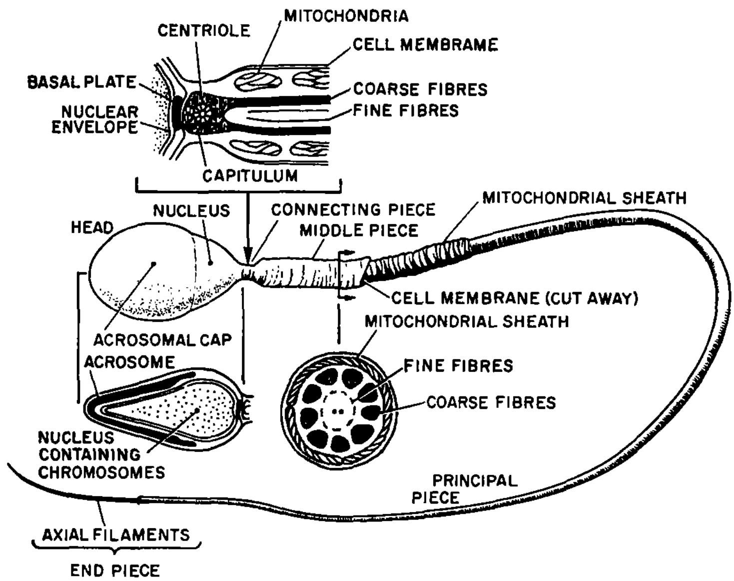


Fig. 126
The human sperm cell
Source: Hancock. 'The sperm cell'. *Science Journal*, June 1970.

second meiotic spindle. Between the cell membrane and a relatively thick transparent zona pellucida is a fluid-filled perivitelline space in which the first polar body is located. Surrounding the zona pellucida is a mass of follicle cells embedded in a gelatinous matrix. An enlarged egg is diagrammed in Fig. 127.

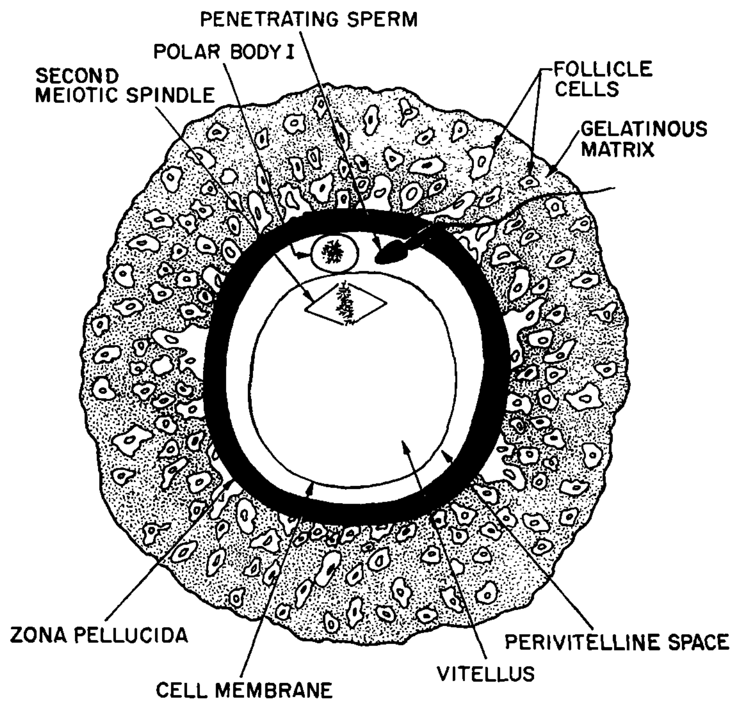


Fig. 127
The human egg just prior to fertilization
Source: Austin. 'The egg and fertilization'. *Science Journal*, June 1970.

When the sperm cell contacts the mass of follicle cells and gelatin, the enzyme from the acrosome enables the sperm to digest a path through to the cell membrane of the vitellus. The sperm then is engulfed by the egg cytoplasm, the sperm head swells, becomes open structured and changes into a nucleus of typical appearance. While these events are occurring, the egg completes its second meiotic division, dispelling the second polar body. Changes in the zona pellucida and the surface of the vitellus tend to prevent the entry of more than one sperm. (When polyspermy does occur it results, in general, in the early death of the embryo.) The climax of fertilization occurs when the (pro)-nuclei of the male and female approach, meet and unite. Fertilization accomplishes a restoration of the diploid numbers of chromosomes through the reassociation of the male and female set; the determination of the sex of the new individual and the activation of the ovum into cellular division or cleavage.

Early development

The fertilized egg, or zygote, undergoes its first cleavage about thirty hours after insemination. Technically it is now an embryo and will not become a foetus until it takes on the recognizable form of a human being. Succeeding mitotic cell divisions occur about every ten hours while the embryo proceeds down the Fallopian tube toward the uterine cavity. About three days after fertilization a tight ball of about sixteen cells, the morula, enters the uterine cavity. Up to this point the embryo has subsisted on the vitellus, that accumulated in the egg while it matured in the ovary. Upon entering the uterus the embryo will depend upon sources of food external to it.

Before implantation into the uterus, the cells of the morula divide and differentiate into a thin-walled vesicle, the blastocyst. The blastocyst consists of two distinct groups of cells, each of which will follow a separate course. An inner cell mass is the precursor of the embryo and several extra embryonic membranes, while an outer shell of cells, the trophoblast, provides an isolating envelope for the embryo and an expanding membrane with which contact is made with maternal tissues. While in the cavity of the uterus, the blastocyst, still no larger than the original egg, obtains its nourishment from tissue fluids which surround it.

About the seventh day after fertilization the blastocyst begins its implantation into the endometrium of the uterus. The rapidly dividing cells of the trophoblast secrete enzymes which erode the cells of the endometrium and enable the penetration of the uterine lining by the blastocyst. Eventually the blastocyst becomes enclosed in the uterus by endometrial healing. At first the blastocyst derives its nutrition from the endometrial tissues digested by the trophoblast. Later nutriments are obtained from maternal blood which bathe the cells of the trophoblast. Eventually the trophoblast, later invaded by the chorion, and the adjacent uterine cells form the placenta, an organ by means of which the developing embryo obtains nutrients and oxygen, and gets rid of metabolic wastes.

Soon after implantation of the blastocyst, the trophoblast begins to secrete the hormone, chorionic gonadotrophin. This substance, acting like LH, prevents the

degeneration of the corpus luteum and assures a continuous flow of progesterone and estrogen. High levels of the latter two hormones ensure the maintenance of the endometrium and suppress the production of FSH and the consequent ripening of more eggs. Eventually (between the fourth and seventh months of pregnancy) the placenta takes over the role of the corpus luteum by secreting sufficient progesterone and estrogen to sustain the pregnancy. However, if the progesterone secretion by the corpus luteum ceases too early, before take-over by the placenta, the pregnancy will terminate in miscarriage.

Implantation amounts to the grafting of a foreign organ to the tissues of the uterus. The physiology of the implantation process is poorly understood and knowledge as to why the mother does not produce an immune response to this foreign tissue should shed light on the problem of rejection associated with organ transplants. Fig. 128 relates events from ovulation to implantation.

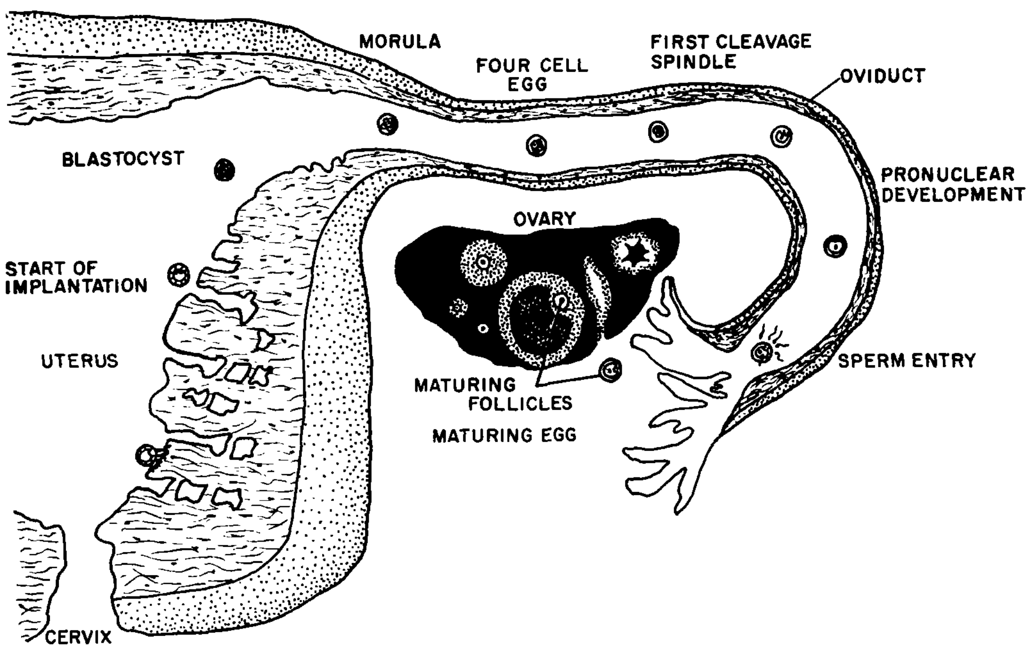


Fig. 128
Sequence of events of fertilization from ovulation to implantation
Source: McLaren and Kerr. 'Egg transplantation'. *Science Journal*, June 1970.

Fig. 129 illustrates the sequence of events from implantation to late embryonic development. The inner cell mass of the blastocyst develops into the human itself and into extraembryonic membranes which nourish and protect the developing organism. The first membrane to form, the amnion, grows around the embryo and forms the amniotic cavity which fills with a watery fluid secreted by the amnion and the embryo. The embryo is thus suspended in a pool of physiological saline which prevents desic-

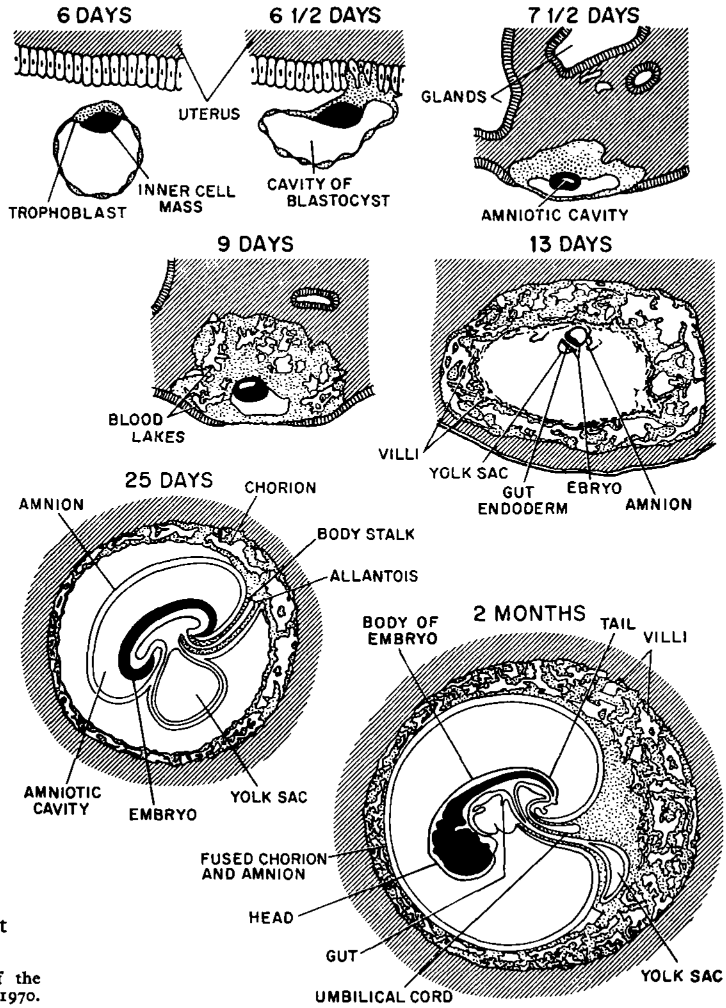


Fig. 129
Early stages in the development of the human embryo
Source: Amoroso. 'Development of the Early Embryo'. *Science Journal*, June 1970.

cation and provides a protective shock-absorbing cushion. A second sheet of cells invades the trophoblast and establishes fingerlike projections, or villi, into the uterine tissue, forming the organ known as the chorion. A sheet of cells grows out from tissues which are destined to become the digestive tract and forms the yolk sac. In the developing eggs of egg-laying vertebrates the yolk sac encloses the embryo-sustaining yolk, but in higher mammals the yolk sac has no known function. Later a second sac, the sausage-shaped allantois, grows out from the gut and fuses with the chorion in the body stalk. The body stalk connects the embryo with the enveloping chorion and enables the embryo to float free in the amniotic cavity. In egg-laying animals the allantois is often a large pouch functioning in the storage of nitrogenous wastes and functioning as a gas exchanger between the embryo and the atmosphere outside the shell of the egg. For higher mammals, tissues of the allantois are the source of blood vessels which serve the placenta.

As the embryo grows the umbilical cord forms at the junction where all the membranes come together and link the embryo with its placental attachment to the uterine wall. The umbilical cord is enclosed by the amnion and includes the umbilical vessels, chorionic tissues and the remains of the yolk sac and the allantois. At birth this cord is 1 cm in diameter and 70 cm long.

As development proceeds, the chorion thickens at the base of the umbilical cord forming a large number of deeply penetrating finger-like chorionic villi which are bathed in pools of maternal blood. This disc-shaped structure, the placenta, enables the passage of nutrients, wastes, gases and antibodies between mother and child. Material exchange occurs normally by diffusion through the membrane of chorionic capillaries. At birth the disc shaped placenta measures 15 to 20 cm in diameter, 2 to 3 cm in thickness, and about 500 g in mass.

The placenta acts as a barrier against some infections of the mother; nevertheless some viruses (measles, influenza, smallpox) may pass through it and infect the foetus. In the third week the embryonic disc begins to undergo rapid growth. At first, thick longitudinal ridges convert the disc into a tubular body. More rapid growth at the anterior end creates a bulging head which folds under, bringing what are to become the mouth and heart to proximity on the ventral side. To a lesser extent the tail underfolds the posterior end of the embryonic disc. By the end of the second month the embryo has developed the principal external features which are distinctly human and the developing organism is now referred to as a foetus. The head becomes erect and the face develops eyes, ears and a nose. With the settling of the heart and the disappearance of the branchial arches and gill clefts, the neck becomes distinguishable. The limbs become recognizable with visible digits. And at the posterior end the tail becomes inconspicuous as the buttocks develop above it. (See Fig. 130 for the position of twins inside the uterus).

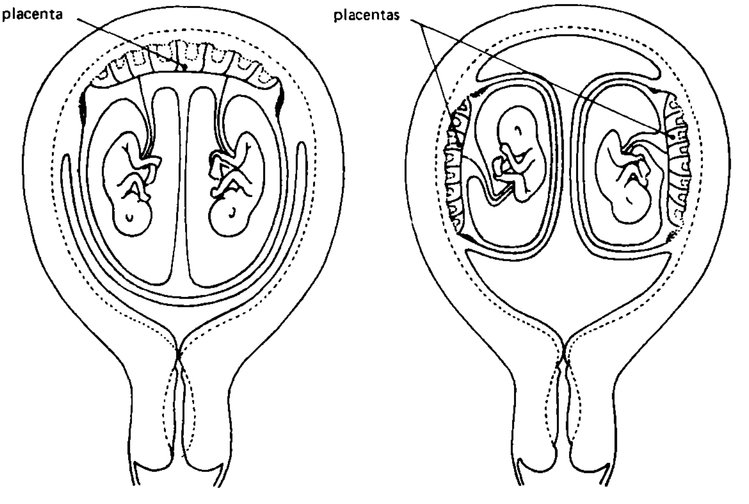


Fig. 130
Position of twins inside uterus

Left: true twins (there is only one placenta to which are attached the two umbilical cords); the two foetuses originated from one egg which divided after fertilization; they have therefore the same genetic make-up.
Right: false twins (there are two placentas and two independent blood circulations); the two foetuses originated from two eggs fertilized at the same time by two spermatozoids; their genetic make-up are as different as those of two non-twin infants born from the same parents.

Birth (parturition)

The human gestation period is normally 40 weeks, counting from the data of the woman's last menses, but live birth can occur as early as 28 weeks or as late as 45 weeks. During gestation the uterus increases to about twenty-four times its normal size and will extend as high as the lower end of the breast bone. The average weight of the foetus by the end of term is 3,000 grammes, or about 1/20th the weight of the mother.

At the end of the gestation period the foetus is usually positioned with its head down and its limbs and head folded in close to the trunk. When labour begins, rhythmic contractions of the uterus cause pressure on the amniotic sac and force it onto the floor of the pelvis. This pressure forces the dilation of the cervix and the ultimate rupture of the amniotic sac releasing the amniotic fluid or the 'waters' (see Fig. 131).

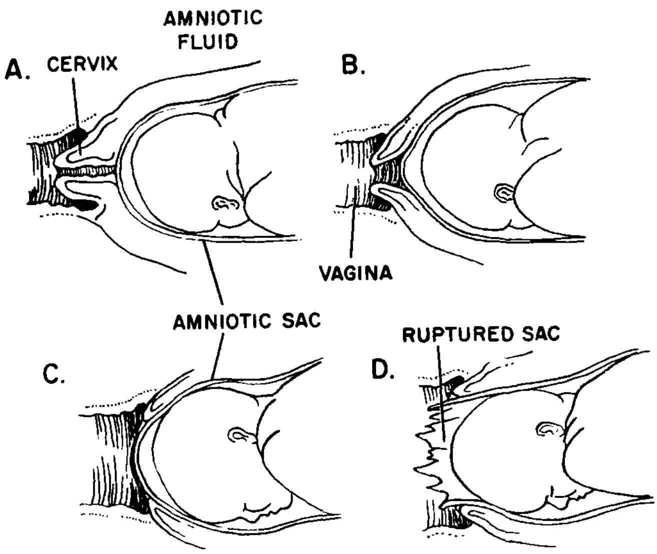


Fig. 131
The first stage of parturition
Source: Goin and Goin. *Man and the Natural World*. Macmillan, 1970.

This first stage of parturition beginning with uterine contractions and culminating with the rupture of the amniotic sac, lasts about twelve hours. The second stage of birth, expulsion, is characterized by strong and frequent voluntary and involuntary uterine contractions which expel the foetus from the body of the mother as illustrated in Fig. 132. Expulsing may require from twenty minutes to one hour to complete. During the final stage of birth, contractions of the uterine muscles result in the dislodgement of the placenta from the wall of the uterus and the subsequent expulsion of the placenta as the afterbirth. Uterine contractions during this stage have the effect of squeezing much of the foetal blood standing in placental blood vessels back into the vessels of the foetus. After the pulsation in the umbilical cord ceases, the cord is tied and then severed. This final stage of birth lasts from ten to fifteen minutes. After the third stage of birth, bleeding occurs from the uterine wall where the placenta

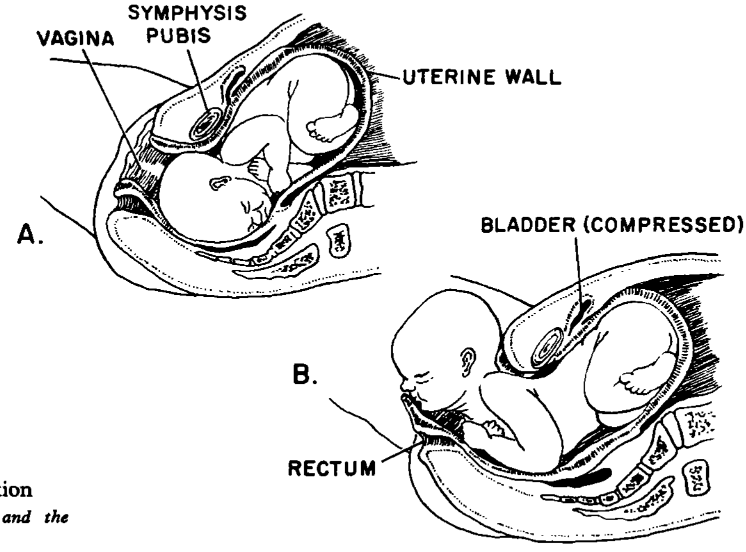


Fig. 132
The second stage of parturition
Source: Goin and Goin. *Man and the Natural World*. Macmillan, 1970.

had made a tight connexion. However, constrictions of the blood vessels of the uterus usually prevent serious hemorrhage.

At birth the placental circulation ceases, blood oxygen levels in the foetus drop, the respiratory centre in the brain is stimulated, the newborn baby takes its first breath, the lungs fill with air, the chest expands, the blood is directed in great quantities to the lungs where it becomes oxygenated.

Through the pregnancy the mammary glands of the mother, under the influence of estrogen and progesterone, enlarge due to the deposition of fat and the engorgement of blood. At birth the hormone prolactin from the anterior pituitary plays a dominant role in stimulating the secretion of milk, but the continuous production of milk depends upon the regular stimulation of the breasts' nipples by the suckling child. The stimulation of sucking causes the secretion of oxytocin. If the infant is not breast fed the breasts will cease to secrete milk after several days and return to their normal size as oxytocin is no longer secreted in the absence of nipple stimulation. Milk production may continue from six months to one and one-half years. There is some evidence that ovulation and menstruation tend to be delayed longer in women who nurse their babies than in women that do not. Normal cycles usually return about six weeks after delivery when the mother does not nurse the child. But women who nurse their children cannot depend upon infertility beyond six weeks postpartum.

17 Sexual behaviour

Introduction

For most infra-human species of animals, sexual behaviour is limited to the act of copulation. The scope of sexual behaviour is extended in the case of such animals as fish and birds, where reproduction involves, in addition to copulation, acts of courtship and the care of eggs and/or youngs. Among humans and some of the lower primates, sexual behaviour includes reproductive performance and those inter-individual pattern of behaviour which involve sexual differences. Male and female primates begin at infancy to play sex roles which establish adult patterns of reproductive behaviour. For example, it is well established from studies of monkeys that the play of infant males and females together is important in the development of male and female postures essential for normal adult sexual behaviour. Similarly, the way in which human children are taught to relate with other children is believed to affect their sexual performance as adults. Human sexual behaviour must be broadly defined to include interactions which permeate the whole fabric of life.

Human sexual behaviour can be examined from three perspectives. The evolutionary view seeks to identify the behavioural elements which are common to human and lower animals, and those behavioural elements which are unique products of human evolution, or are modifications of behaviour which societies impose on the genetically directed tendencies of its members. A second perspective examines the relationships between behaviour and the physiological condition of individuals. The chemistry of the blood and the structural and functional capacities of the nervous and muscular systems determine kinds of behaviour and explain many of the differences and similarities among humans, and between humans and other species. The third perspective is provided by cross cultural evidence. The sexual behaviour of no one human society is representative of the behaviour of the human species. By objectively analysing sexual behaviour in many human societies it may be possible to describe comprehensively behaviour as it is modified by culture, and identify behavioural patterns which are common to all human societies either as a result of a biological heritage or learning experiences common to individuals of all cultures.

In the following sections, specific sexual behaviour patterns are described and

discussed from evolutionary, physiological and cross cultural perspectives. The topics are sequenced to proceed generally through the stages of sexual development.

Development of sexual behaviour

Cross-cultural and cross-species evidence indicates that patterns of sexual behaviour are neither wholly genetically established in the individual by birth nor are they entirely socially determined. The development of sexual behaviour is an interaction of learned cultural values, experiences, physiological drives, tendencies, and capacities.

Newborn infants have neuromuscular systems which when stimulated carry out sexual acts. The penes of infant males when stroked will engorge with blood and become erect. Continued stimulation leads to reflexive copulatory thrusts and orgasms. Copulatory responses can be induced in infant females by stimulation of the clitoris.

Functional sexual behaviour does not commence until late in childhood when the anterior lobe of the pituitary gland begins the heavy secretion of the gonadotrophic hormones. These hormones from the pituitary stimulate growth and activity in the ovaries and testes. The gonads respond first by secreting their own hormones and then after about a year they begin producing mature eggs and sperm. In the male, the physical effects of the hormones from the gonads include the development of pubic, underarm and body hair, the deepening of the voice, the broadening of the shoulders, and the ability to produce semen. In girls the enlargement of breasts, the development of pubic and underarm hair, the broadening of the hips and the onset of menstruation (menarche) signal the beginning of sexual maturity, the stage of development known as puberty. The period between puberty and full psycho-sexual maturity is known as adolescence. Puberty may begin in girls around the ages of 11 and 12 years, and about a year later for boys. There is considerably greater variation in ages of girls and boys for the onset of puberty within populations, than among populations. The major determining variable, aside from genetics, appears to be diet. The duration of adolescence is also variable lasting sometimes in excess of ten years.

Physical sexual precocity is often caused by the untimely secretion of abnormally heavy amounts of sex hormones quite frequently stimulated by tumours in endocrine glands. Sexual precocity in males sometimes accompanies the hormonal treatment for cryptorchism to induce the descension of testes. Sexual precocity, due to untimely hormonal flow is not a cause of mental retardation, but interestingly it does not ordinarily render the child precocious with regard to sexual behaviour.

Studies of the behaviour of juvenile monkeys and apes are particularly significant in developing theories about the role of the interpersonal environment of the immature primate in establishing its adult patterns of sexual behaviour. When chimpanzees are reared apart and then brought together in pairs when the female is in physiological heat, they fail to achieve successful sexual unions. The female will present herself to the male, and the male may develop an erection and make awkward attempts to mount her, but they make too many wrong moves in their failure. Monkeys reared with surrogate mothers and without contact with other infants also fail as adults to achieve

successfully penile insertion, even though they react to sexual stimuli with recognizable positions of the coital pattern. The female may present herself too high, or the male's thrusts may be temporally or specially misdirected. They appear to have all the necessary reflexes, but as often as they try, they fail to copulate. They appear to lack the necessary knowledge of how to mate. If one of the chimpanzees or one of the monkeys has had previous successful sexual experience, successful copulation will often occur. If the female is an inexperienced partner reared in the absence of other members of the species, she will abuse and reject offspring that result.

Studies of infra-human primates reared together have shown that infants, children and adolescents engage in extensive sexual play. In the course of their development they have been observed to reproduce all the sexual behaviours of normal adults. It is believed that the spontaneous sex play of immature infra-human primates develops the competence in these individuals to engage successfully in heterosexual coitus by the time they become adults.

Cross-cultural evidence suggests that when human children are raised in sexually permissive societies they engage in kinds of sex play similar to those observed among juveniles of other primate species. Further, the evidence continues to suggest that these individuals with unrestricted knowledge and experience in heterosexual relations, reach better sexual adjustment in adult life. On the other hand, children raised in societies which repress the sexual tendencies of children to play with their own and other's genitalia, to enjoy sensual physical contacts with other children, to observe and imitate the sexual behaviour of adults, have less opportunities to achieve satisfactory sexual adjustments in adulthood.

Attraction for a sex partner

The physical appearance of males and females plays a significant role in attracting sex partners among humans. In cross-cultural perspective there are few criteria of beauty. In general, extreme obesity, poor complexion, baldness, masculinity in females and feminity in males, and advanced age are sexually unattractive characteristics in human societies. Usually the traits of females are considered more important than the traits of males. The sex appeal of a woman would change considerably as she would be viewed by the men of various cultures. Where men of one culture prefer small round breasts, men of another favour large breasts, while long and pendulous breasts are admired elsewhere. Similar variations of preferred traits include factors relating to body hair, eye shape, skin colour, height, body form, and size of the external genitalia. The standards of beauty are so varied among human societies that international beauty contests are impossible to judge fairly.

Aside from physical features people employ odours (perfumes), sounds (music), gifts, dances, colourful clothing and ornaments and such abstract symbolism as poetry and love letters to attract sex partners. The behaviour of human lovers *appear* to have a counterpart among infra-human animals which employ audio, chemical and visual sexual attractants. But one must remember that these human behaviours employ

language, and are used in societies where direct solicitation for sexual intercourse does not occur. Infra-human animals communicate their physiological readiness for copulation when the body chemistry directs the performance of mating calls, the mating dance, or the pursuit of an odorous mate.

Among mammals the exposure of the female genitalia, as when the female turns her back to the male, raises her tail and raises her back, is the final and compelling attraction to the male. In human societies, the exposure of the female genitalia invites sexual assault by males. Quite likely, it is for this reason that sexually mature females are taught in practically all cultures to cover their genitalia with clothing.

Sexual partnership

The large majority of adults in all human societies are married. The form of marriage differs among societies and may take on one of four possible forms: monogamy, one male and one female; polygyny, one male and more than one female; polyandry, one female and more than one male; and polygamy or group mateships, two or more males and two or more females. Despite the fact that most human societies permit one mate (usually the male) to maintain more than one sex partner, the monogamous marriage is becoming the prevalent form of sexual partnership throughout the world, economics, social status and religious belief tending to be limiting factors.

No human society condones indiscriminate mating. To guard against unstable liaisons every society directs and restricts the selection of sexual partners and thus ensures the establishment of fairly stable partnerships. Societies do vary, however, in their permissiveness toward extramarital relations.

Most infra-human primates establish fairly permanent polygamous mateships. Some lower animals also maintain stable and enduring sexual partnerships as has been observed among wild species of Canidae.

Every known human society prohibits incest—i.e. sexual relations or marriage between father and daughter, mother and son, or brother and sister. The definition of prohibited incest may often be enlarged to include some members of the extended family. Animals in the wild with both family and wider social groups, also appear to prohibit against incest within the nuclear family though the behaviour may manifest as aggressive sexual jealousy. Domesticated animals, on the other hand, frequently engage in incest.

Among the probable reasons for the universal, intensive and ancient restrictions against incest among humans are: incestuous liaisons often produce genetically inferior offspring; incestuous love affairs within families create divisive rivalries, passions and jealousies; families must establish alliances through marriage with the social group for reasons of mutual assistance; and, in order to rear children successfully the parental roles must be separated from a sexual role and the taboo against incest is taught early and severely.

Sexual stimulation

Sexual play prior to intercourse serves several biological functions. Among infra-human animals with courtship ceremonies, the routines of males and females are specific and tend to enable potential sex partners to pass a test of identification, thus ensuring an intraspecies mating. For all animals, including man, foreplay raises the level of sexual excitement and increases the probability that coitus will ensue. The excitement resulting from foreplay also synchronizes the reactions of couples and enables the successful completion of coitus when attempted.

One of the most obvious differences between humans and other animals is the way which males and females initiate sexual interaction. Through language a human will consciously or intentionally attempt to arouse a potential partner, and thus prepare the way for physical contact and greater sexual excitement. Animals without the intellectual powers to interact symbolically depend upon direct physical contact for the arousal of sexual excitement. They are brought into physical contact without the benefit of the knowledge of what will result from their contact. For example, the dog is initially attracted by the odour of secretions of the external genitalia of a bitch on heat. As he licks the vulva he excites the bitch and he excites himself. Their ensuing act of copulation is unpremeditated.

The most common form of bodily stimulation prior to coitus among humans is the handling or mouthing of erogenic zones of the partner. Some of the erogenic zones are the hips, breasts, genitalia and thighs. Among primates and many human societies the grooming of the partner's body is commonly practised. Mouth-to-mouth kissing and the fondling of the female's breasts appear to be a nearly exclusive human practice, but not all human societies practice these forms of sexual foreplay. The mutual infliction of pain (biting, scratching, striking) in association with sexual stimulation occurs within many human societies. The effects of these assaultive practices are to heighten the state of sexual excitement. The abnormal extension of this behaviour is sadism.

Women may develop a dislike, or fear, for sexual interaction and consequently become sexually unresponsive or frigid. One cause of frigidity is a history of failure to become adequately or sufficiently aroused prior to intercourse. Having never derived pleasures from sexual intercourse, it is understandable that the female would be non-responsive to the advances of a male. The once widespread practice of excision (female circumcision) results in the removal of the clitoris and the creation of a condition of frigidity. Presumably the objective of the operation had been to deprive the female of the pleasures of sexual intercourse so that they will regard sex as a pleasureless child-bearing task. A second objective was to better assure the fidelity of wives who, without clitoris, would not be easily attracted into extramarital liaisons.

The sexual act

In all infra-human mammals the mating pattern is essentially four legged with the male entering the female from the rear. Human coitus occurs while the male and female face each other, but a variation of this basic pattern (ventral-ventral) occurs according to habits and culture of different societies. The change in pattern from rear entry to face-to-face entry reflects the ventral shift of the vaginal opening which accompanied the rotation of the pelvic skeleton as the human evolved into an erect posture.

One consequence of the unique face-to-face position of human coitus has been to enable the female to experience orgasm—an explosive, temporarily exhausting, release of built-up excitement—comparable to the release the male experiences upon ejaculation. In large measure female orgasm is explained by the intense stimulation of the clitoris during intercourse while the clitoris is pressed between the pubis bones of the pair and massaged by the strokes of the erect penis. The variations of positions of face-to-face coitus determine the degree to which the clitoris is stimulated and the frequency of the female orgasm. Many human cultures prescribe or subtly determine coital positions and consequently influence the responsive role of the female. Clitorial stimulation is but one determinant of female orgasm. While women appear to have the physiological capacity for orgasm, its achievement in women of some cultures may be dependent upon the psychosocial acceptance of their sexuality. In certain human cultures the female is conditioned to assume a passive role during coitus precluding the activity and attitude prerequisite to the achievement of orgasmic satisfaction. In general it is considered that a fair sexual adjustment and a reciprocal satisfaction and love expression may be achieved in a human couple when simultaneous orgasm is reached by both partners. This simultaneous orgams can be achieved more often when both partners are adequately sexually excited. There is no evidence to suggest that any infra-human female animal achieves orgasm during coitus.

The duration of human coitus varies among cultures and among individuals. The controlling factor is the speed with which the male achieves orgasm. The males of some cultures have learned to control the muscles which initiate ejaculation, and can reach orgasm at will. Studies of coital duration in other human cultures reveal ranges from less than a minute to longer than thirty minutes between the time of intromission and ejaculation. Coitus among apes lasts between ten and fifteen seconds, and the duration for most other mammalian species is also brief. On the other hand, exceptions such as mink, alpaca and sable have been observed to remain in the coital embrace for as long as eight hours.

While coitus between males and females is prevalent in human and all mammalian species, other forms of sexual release are commonly observed. Human practices of homosexuality, masturbation and animal intercourse are known to occur in all cultures despite restrictions. Instances of non-heterosexual coitus are less frequent among infra-humans than humans. Young apes and monkeys indulge in frequent noncopulatory acts, but upon maturity male-female coitus prevails.

In general, humans of most societies copulate at night. Since humans are diurnal, and animals tend to copulate during the portion of the day during which they are most

active, the human tendency may reflect the desire for privacy and the learned satisfactions of associating copulation with sleep. The places where humans copulate appear to relate again with the availability of privacy. In societies where couples cannot easily engage in copulation in the house without detection, they will characteristically copulate outside the dwelling.

Although a human couple in the full bloom of life may have the physiological capacity to engage in sexual intercourse regularly one to four times daily there are numerous social rules which interfere with capacity performance. For many societies marriage is a control against intercourse outside the institution of the family. Within marriage intercourse is further controlled within various societies by restrictions that demand abstinence during pregnancy, menstruation, period of post-natal recovery, mourning, and religious holidays.

Other sexual activities

Universally humans are known to manipulate their sex organs in ways that produce sexual excitement. For most males and many females these acts of self stimulation are carried out to orgasm. Although many observations are of animals maintained in captivity, masturbation among apes and monkeys is well known. The above facts suggest that human masturbation is a normal and natural practice although the one who does not practice masturbation cannot be considered abnormal. The fact that the occurrence of self stimulation is less common in societies which rule against it than in societies that do not, is another example of the power of social conditioning. Among all mammals, masturbation has lower frequency in females than in males. Some humans are also capable of sexual stimulation, and even orgasm, without physical stimulation, by sexual fantasy. The male's nocturnal emission may possibly occur without the physical stimulation of genitalia. Such spontaneous orgasms have also been observed in infraprimate animals.

There is a tendency in mammals, including man, toward sexual activities between members of the same sex. Homosexuality, or sexual inversion, occurs in all human societies, but is never the predominant form of sexual behaviour. The incidence of homosexuality is affected by the degree of social restriction, being less common in societies where children are strongly conditioned against it and toward heterosexuality. Homosexual behaviour is more common in adolescence than in adulthood, and is practised by more men than women.

The most common methods of sexual release among homosexual males are mutual manual masturbation, stimulation of the penis with the lips and tongue (fellatio), anal intercourse (pederastry), and femoral copulation where the penis is inserted between the thighs. Among homosexual women the methods of stimulation include hugging and kissing with considerable body contact, the mutual manual or oral stimulation of either self or partner's clitoris, usually to the achievement of orgasm. Rarely do homosexual women employ a penis substitute which is inserted into the vagina.

Although sexual relation between humans and animals (bestiality) does not occur as frequently as homosexuality it is common enough within human societies to warrant mention. The acceptance and occurrence of bestiality vary considerably among human societies and reflect the ability of a society to channel the sexual behaviour of its people. Nevertheless the human, especially the adolescent male, is capable of being aroused and achieving sexual satisfaction through *inter alia* vaginal or anal intercourse with an infra-human animal.

Hormones and sexual behaviour

The sexual behaviours of lower animals, birds, and fish are strongly controlled by gonadal hormones. In the female the level of estrogen in the blood increases as the breeding season comes on. Under the influence of the hormone the female will seek a partner, copulate, and care for young. The power of the hormones to direct behaviour can be demonstrated by inducing sexual responsiveness out of season through the administration of hormones, or conversely, terminating sexual behaviour by removing the hormone-producing gonads (castration).

The effects of learning on the sexual behaviour of infraprimate animals are minimal. Some sexually experienced male mammals have been observed to copulate after castration and even long after testosterone levels have dropped considerably. Full performance can usually be restored with hormone treatment. Inexperienced male castrates, however, do not perform sexually even after hormone therapy. Learning and experience apparently play a role in the sexual behaviour of these male animals. However, the infraprimate female is very much a behavioural prisoner of her sex hormones.

Female monkeys and apes, like humans, maintain continuous fertility cycles

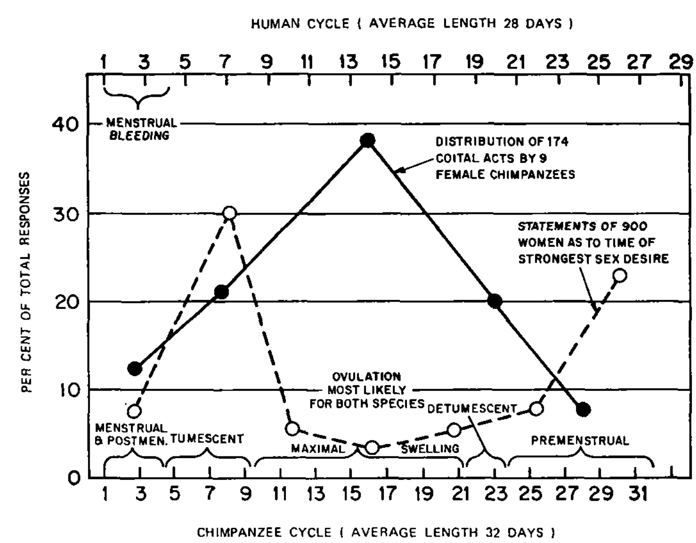


Fig. 133
Cycles of sexual desire described by 900 American women compared with cycles of coital behaviour shown by 9 female chimpanzees
Source: Ford and Beach. *Patterns of sexual behaviour*. Harper, 1951.

throughout their reproductive lives. Unlike lower animals, they do not experience seasonal breeding. Further, the female's reproductive capacity and sexual responsiveness are geared to a hormonally controlled menstrual cycle. The male infra-human primate, like many male mammals, is capable of sexual performance any time throughout his adult life.

One difference between the sexual behaviour of infra-human primates and humans is seen in the relationships between the fertility cycle and the apparent rhythm of sexual desire in females. Fig. 133 shows that in 174 matings between chimpanzees, by far the greatest number occurred during mid-cycle—when genital swelling was maximal, estrogen levels highest, ovulation imminent and pregnancy probability greatest. Contrasting data for people relate the female desire for copulation with the menstrual cycle, since frequency studies of human copulation reflect the continual noncyclic desires of man who in most societies decide when coitus will occur. From a sample of 900 women, the times of sexual desire peaked immediately before and after menstrual flow. One explanation for this phenomenon notes that almost every human society prohibits coitus during menstruation. Pre- and post-menstrual peaks of sexual desire may be the effects of deprivation of intercourse during the period of menstrual abstinence.

The human female is less dependent upon hormones for sexual responsiveness than her primate relatives. Further evidence of her greater freedom from hormonal dependence is seen in the behaviour of menopausal women who are beyond the age of reproductive capability, and those women who have experienced surgical menopause through the removal of both ovaries. Menopause is characterized by a fall in estrogen levels and a rise in gonadotrophins. The female suffers a variety of discomforts as results of these hormonal changes, but women do not suffer, however, as a consequence of the loss of functioning ovaries any reduction in their ability to respond sexually. On the other hand, infra-human female animals cease their sexual responsiveness upon the fall of hormone levels. The sexual behaviour of the human female appears to be far more strongly influenced by learning and experience than by hormones, and this fact is another illustration of how in the evolution of humans the centre of sexual behavioural control shifted from the gonadal glands to the cerebral cortex of the brain.

Men are much less dependent upon gonadal hormones for sexual performance than are lower male animals. Castration of the sexually experienced human male need not affect his normal ability to become sexually excited and perform the sex act to orgasm. Castrated sexually experienced monkeys and apes often continue to copulate for several years, while lower animals show a relatively sudden decline in sexual responsiveness after the loss of the testes. Normal sexual behaviour can usually be restored in castrated and aged males through the administration of hormones but less markedly as one proceeds up the phylogenetic tree. In the human male, hormonal therapy is often not necessary for some castrates, effective for others, and ineffective in restoring or providing virility to others. Apparently, nonhormonal factors play major roles in regulating the sexuality of men to almost the same degree that the experience and social conditioning of women, more than anything else, direct their sexual behaviour.

Further reading

FORD, Clellan S. and BEACH, Frank A. *Patterns of sexual behaviour.* New York, Harper and Brothers, 1951.
HUMAN REPRODUCTION. *Science Journal* (London), 6, no. 6, June 1970.
KEYFITZ, Nathan and FLIEGER, Wilhelm. *Population: facts and methods of demography.* San Francisco, Freeman & Co., 1971.
LORIMER, Frank. *Culture and human fertility.* Paris, Unesco, 1954; 2nd ed. 1958, 514 p.
MASTERS, William H. and JOHNSON, Virginia E. *Human sexual response.* Boston, Little, Brown & Co., 1966.
UNESCO. Projet pilote pour l'enseignement de la Biologie en Afrique. *Quelques aspects de la biologie humaine* (reproduction). Paris, AUDECAM, 1972, 36 p.

Part V

Present and future evolution or Design for survival

18 Ecosystem management and control of environmental quality

ECOSYSTEM MANAGEMENT

Introduction and definitions

Man has altered the biosphere in a relatively short time to such an extent that one can no longer divide the environment into 'human' or 'natural'. Very few ecosystems at the present time are so uninfluenced by human activities that they can be labelled 'natural'. Air- and water-borne pollution is affecting all seas and continents, the Antarctic and Arctic polar caps not excluded. Pollution is just one example of how man affects ecosystems. Yet ecosystems and environments are still functioning in the same way as in remote times. There is no technological substitute for this function.

Man's environment is a closed system, where life goes on through conversion cycles, transfer of organic matter, movements of air, water and soils. In these processes all the renewable natural resources interact as components of a functional system. Hence human as well as animal populations are integrated parts of these components. They all work together like cells and organs in a living body. The degradation of any of these resources has effects on all the others. Ecology is the basic science on which maintenance, conservation, management and utilization of natural resources, as well as the welfare of man, depend. In fact, ecology comes very close to bioeconomy. Ecosystem productivity and conversion cycles are processes which could be expressed in economic terms. This is a useful bridge of communication between ecologists and economists enabling the latter to realize that the ecosystem is a production system with a balance between losses and replacement. This view must also underlie political decisions.

An intelligent use of renewable natural resources implies reaching an ecological balance between man's needs and nature's long-term capacity to satisfy them. In other words, conservation and management of renewable natural resources include their utilization on a sustained yield basis, that is without deterioration. It is evident that certain of the renewable natural resources are deteriorating mainly as a result of unwise land-use practices which have been going on for a long time. In addition, new forms of environmental danger have recently become apparent in many developing countries, namely the pollution of the air, water and soil, often through the use of toxic chemical pesticides which are accumulating on an accelerating scale in the world environment and all its living organisms, including man.

The central question is: how, where, and when to harvest biological resources? To exploit them wisely does not mean only a direct economic profit. The manifold riches of nature are not simply economic, they are also social, cultural, scientific and aesthetic.

Objectives for the management of ecosystems will vary from one region to another depending on a number of factors and the needs of the various countries. Nevertheless, to be successful it must always be based on sound, long-term ecological research of an interdisciplinary character. This will enable land-use planners to understand the structural and functional ecology as well as the successional evolution at ecosystem level of the area concerned, as a basic requirement for ecological considerations preceding decisions on management objectives.

An important feature of tropical and subtropical ecosystems is their vulnerability to misuse in relation to the case in temperate areas. This is due to many factors. The fertile soil layer is usually thin because of a high ecosystem metabolism and energy flow in which decaying organisms break down organic matter very quickly to be immediately re-used as nutrients through recycling. This is particularly pronounced in forest ecosystems but also in other types of biomes. Therefore, if the vegetation cover is removed in an area where the climate is characterized by hot and dry periods interspersed with violent rainy seasons, as is usually the case in Asia, the soil becomes exposed to sun, winds and rains leading to serious erosion, a rapid disappearance of humus and an accelerating rate of soil loss.

Specific conservation measures

When describing 'specific conservation measures' for a whole continent, it is obviously necessary to speak in general terms. The word 'specific' refers here to various measures which must be undertaken in order to manage each one of the different renewable natural resources in a proper way without forgetting that they are all interacting. The aim of conservation measures should be to maintain high productivity of an area or to restore an area in order to regain its ecological potential and diversity. It is also necessary to recognize and foresee the needs of future generations.

Policy, research, conservation and management are essential for the wise utilization of renewable natural resources. All these components are dependent upon each other. Therefore, conservation problems must be tackled in an integrated way. Research data must be available prior to policy formulation. When the land-use objectives of an area or a region have been determined, monitoring and research must go on in various fields as a basis for management. But management may also set priorities for research. Finally, both research and management may induce policy revisions.

Air resources

Large parts of Asia seem still to be relatively unharmed by air pollution but it is becoming increasingly evident that the problem exists. However, no precise data are available due to the lack of monitoring. (See the section on 'control of environmental quality' in this chapter.)

Marine resources

The marine ecosystem was dealt with in Chapter 9 on 'the oceans'. Some urgent conservation measures to save and restore the productivity of the sea resources are the following.

1. Total prohibition, backed by strict control of the use of explosives as a method of fishing.
2. Protection, backed by strict control, of coral reefs.
3. Strict regulation as regards collecting live marine invertebrates for the commercial curio trade.
4. Increased surveillance of oil pollution from ships.
5. Increased surveillance and monitoring of marine pollution from land-based factories and plants as well as from municipal waste discharged into the sea.
6. International agreement regulating industrial fishing from factory ships off the coasts.
7. Marine turtle restoration programme.
8. Mangrove vegetation is a productive marine resource which needs to be utilized properly.

Freshwater resources

Some measures to be taken are the following:

1. Conservation and management of watersheds.
2. Postponement and/or revision of planned and current water development projects in arid regions until investigations have produced a clear picture of the long-term ecological consequences.
3. No new irrigation, drainage and hydro-electric schemes should be undertaken unless they are preceded by ecological investigations and unless it is determined that the long-term consequences of these projects are not adverse to the general productivity of the region concerned and to the welfare of the inhabitants utilizing the region.

4. Monitoring of water pollution in rivers, lakes and groundwater.
5. Investigations of the fertilizing effects of flooding on the soils of certain river valleys and flood plains.
6. Investigations of the productivity of permanent and seasonal wetlands.

Soil resources

Many means for halting erosion and improving soils are at hand. Most important and urgent is restoration of the vegetation cover in order to stop the erosive actions of wind and water. The next step is to let grass and forests grow again on lands that have become too arid and poor to support cultivated plant life. Reforestation will also help to restore the local climate. Another way is to utilize the water resources properly. Mechanical and engineering measures such as terracing, water reservoirs, and the like may be necessary additional aids to restoring the soil, but cannot alone achieve positive results. They must be preceded or accompanied by basic biological measures.

The ultimate goal of soil restoration and conservation is to put the biological processes of soil formation and soil life cycles back to work as they functioned before the vegetation was destroyed. This is a long-term task. It may well take 500 to 1,000 years or more to restore forests that once grew on mountain slopes. These forests, as well as lowland forests and grasslands, over the course of millennia had themselves created the soil humus on which they based their existence. And before the trees colonized the land, other types of vegetation—also during thousands of years—had prepared the ground for successive plant communities to evolve, building up to and producing a maximum conversion rate or energy flow in which the soil was an imperative force.

Plant resources

In heavily grazed and trampled areas the reduction of vegetation cover becomes so great that the soil is laid bare and erosion follows. Very numerous semi-desert and desert areas in Africa and Asia are the end products of the negative evolution which has taken place in only a few decades. Range management for livestock production in the tropics and subtropics is not an easy task. Many agriculturists and veterinarians believe the solution is simply one of reducing livestock numbers, but it is much more complex than that. As noted, cattle, goats, and sheep do not behave like wild ungulates. They keep closely together. Grazing, drinking, licking salt, walking, ruminating or sleeping, they never spread out over the grasslands—a habit that simplifies the herdsman's job. They literally wear away the vegetation on some areas of the range before moving to another, topographically less favourable site.

Therefore, it is important to initiate a long-term restoration programme coupled with other measures as follows:

1. A long-term restoration programme in order to regenerate the grass cover, the woodlands and the forests, which will contribute to the restoration of the water, soil and wild animal resources.
2. An ecological survey and assessment of the productive potential of main vegetation types leading to a land capability map as part of general restoration plans.
3. Protection of the still-existing forests is indispensable, especially in montane areas, where the most important role of forests is an ecological one as maintainers of the water régime and soil fertility. Strict control of fires and controlled grazing must be part of this protection.
4. Regulation of wood-cutting for domestic fuel as well as of charcoal burning and export.

Wild animal resources

The wild animals have evolved on the grasslands and in the forests and are in harmony with the vegetation, climate, water resources and the soil. Therefore, this complex of resources is often a much more productive unit on a sustained yield basis than lands utilized by livestock, particularly on marginal lands.

Whether it involves wildlife management or animal husbandry, or both, it is essential to adjust animal populations to the carrying capacity of the habitat. A high production of proteins and hides can be constantly maintained from wild mammals on lands that could deteriorate under other forms of use.

Wildlife management includes restoring, protecting, conserving and maintaining animal populations. Some measures to be taken are the following:

1. Establishment of buffer zones around all game and forest-game reserves as well as national parks. In these zones, hunting should always be controlled and regulated (in time and space) with adjustments to the status of the animal populations.
2. Moratorium on hunting for a number of years on a regional basis in order to allow the recovering of animal populations to levels which later can be harvested on a sustained yield basis.
3. Creation or revision of legislation acts fitting a modern approach to wildlife conservation, management and utilization including close seasons for vertebrates, so that they cannot be hunted during reproduction periods.

Land use and planning

Since all the renewable natural resources are interacting, solutions to environmental problems, connected with these resources, require an integrated approach. Likewise, land-use planning and a land-use policy must necessarily be based on the totality of renewable natural resources as a whole unit. In the past, planning, policies, management and utilization of renewable natural resources have almost exclusively been organized along sectoral lines, because ecological considerations were mostly absent.

Therefore, an important part of development plans should include a clear-cut national policy on water and land use including the fullest integration of air, water, soil, wild vegetational and wild animal resources into the over-all planning and utilization of a nation's resources. Such an integration must, of course, be based on the real economic and ecological values and potentials of these renewable natural resources in comparison with other resources (agriculture, livestock, etc.), and their effects on the environment.

Such knowledge can only be obtained through ecological surveys of the land leading to the compilation of land capability maps of each region. This material would be the basis for an interdisciplinary planning of various forms of land use in each country or region.

Since renewable natural resources are dynamic and changing, particularly when utilized by man, such ecological surveys covering the whole of a country have to be undertaken at regular intervals in order to provide the government with the necessary background data for a continued, rolling planning process.

Ecological land and water surveys necessarily involve team work. A specialist group, ideally, should include a climatologist, geomorphologist, a limnologist, a plant ecologist and an animal ecologist. It is useful also for such integrated surveys to have sociologists and anthropologists on the team. Interdisciplinary land surveys have been made with important results in a number of countries.

It is of fundamental importance both for determining a policy and drawing up plans for long-term research to know the past history of the area concerned and to what extent it has been influenced by human impact. Each land or water system is the product of a long history of landscape formation. Without facts about the historic background to the present ecological setting of an area, there is the risk of making serious mistakes in both research and management, as well as in establishing a policy, because factors determining the prevailing situation are not understood.

The basic question behind each land-use survey and planning is to find the ecological potential of the region concerned. It can be expressed in many ways. For management purposes, whatever the final goal of a land development plan, the biological productivity of a given area, as a part of the energy flowing through the ecosystem, gives a firm basis of practical information. Among many data emerging from ecological land surveys is the carrying capacity of an area. It gives a measure of the quantity of individuals of any domestic or natural species of plants and animals that the area concerned can support. However, it must be borne in mind that the carrying capacity is not fixed. It varies in many ways in interaction with several environmental factors, and is of course greatly influenced by human action.

The expected answers resulting from any land-use survey, besides the ecological potential of an area, would include the ecological feasibility for a planned type of land use. When this is determined, other queries such as social feasibility and economic viability of a project can be answered.

A useful and desirable end product of a land-use survey is a land capability map. Such a map is based not only on physical factors of the environment (topography, geology, climate, water regime, soils, vegetation, fauna, human population density) but also on social factors and other features of human activities including land-use history. Obviously, aerial photographs are particularly useful as a basis to land capability maps. In fact, the cost of modern land surveys can be considerably reduced by interpreting from aerial photographs.

National parks and equivalent reserves are essential components in any system of land-use surveys and land-use planning, not only because they constitute a wise land use in the form of recreation, education and research, but also as sample areas for comparison with regions which have been modified by man. The important role of national parks and equivalent reserves in the fields of international conservation, research, education and recreation as well as national economy and development is becoming increasingly recognized in many countries on all continents and also by international organizations which are not directly involved in conservation activities. So far, however, national parks have been primarily regarded as national assets. The traditional way of selecting areas as national parks has mainly been based on special features of a country such as spectacular landscape scenery, or rare and interesting geological sites, vegetation and/or fauna. In fact, the areas of many national parks have been chosen to preserve natural curiosities. Also areas where historical events have taken place have in certain countries been set aside as national parks. It is not wrong to use such criteria as national values for selection of national parks, but they emphasize single features rather than the natural scene as a whole (habitats, biomes, ecosystems) and such a limited approach may in the long run be dangerous, because it might lead to unexpected surprises and the disappearance of what one wanted to preserve.

One may ask whether present and future conflicts of land use, due to population and economic pressures or political aspirations, can be reconciled with the idea of having large areas set aside for national parks. These are often regarded as being unproductive. This is a wrong view. It is indeed as vital to preserve habitats and ecosystems as it is to set aside areas for other human needs. This is just the essence of ecological planning on a continental scale.

National parks also belong to the few areas of the world where long-term investigations in quality and quantity, biomasses, energy turnover and conversion rates, evolution and speciation and so forth, can go on without human intervention. We need such comparisons from which governments, landscape planners and, in fact, society as a whole, can draw conclusions and lessons.

CONTROL OF ENVIRONMENTAL QUALITY

Climate modification and control

Climate is a complex of several basic environmental factors. The best known of these factors are temperature and moisture, with the latter generally expressed as relative humidity. This is the amount of moisture in the air at a given temperature as compared with the amount of moisture the air can contain at that temperature, expressed as a percentage. Other factors sometimes considered part of the climate complex are atmospheric pressure, wind, various types of radiation, and air pollution. The prevailing conditions caused by these factors over a period of years in a given region constitute the climate. The conditions caused by these factors at a given point in time constitute the weather. Prolonged and extensive weather modification thus results in climate modification.

In Chapter 5 on 'atmosphere' in Part II, evidence is presented that the carbon dioxide concentration of the atmosphere is currently increasing due to the activities of man. Since carbon dioxide effectively absorbs infra-red radiation, an increased concentration of atmospheric carbon dioxide results in greater atmospheric heat retention. It is therefore expected that a world-wide warming trend would accompany the increasing atmospheric content of carbon dioxide. Such as gradual warming trend was indeed observed from about 1890 to the 1940s, and a mean global temperature increase of about 1.6° C has been realized. However, since the 1940s, there has been a cooling trend with a drop in mean global temperature of 0.3° C. This is in spite of continued increase in atmospheric carbon dioxide, and the explanation for this phenomenon is not certain. One interpretation is that a cycling cooling process has occurred. This hypothesis seems quite plausible since substantial world climatic changes occurred before the industrial revolution. A second hypothesis ascribes these changes to a decrease in solar energy reaching the lower atmosphere due to increased air pollution and atmospheric haziness from industrial particulate emissions, domestic fuel combustion and dust from mechanized agriculture, earth moving and transportation. Vapour trails from high altitude jet aircraft may also be contributing to this change.

The source of greatest concern on world climate modification by man is our

inadequate information. We do not know the full effects of the changes we initiate (a substantial warming trend would result in increased melting of Arctic and Antarctic ice caps, substantial raising of sea level, flooding of many of the world's major inhabited areas).

The climate of cities

Table 89 provides a summary of some major climatic features in cities and compares them with rural surroundings.

Solar radiation striking horizontal surfaces in the city is reduced by particulate air pollutants. The over-all reduction is often as high as 15 per cent though it varies across the radiation spectrum and is usually greatest at ultra-violet wave-lengths.

Table 89
Climatic changes produced by cities

Elements	Comparisons with rural environments
Contaminants	
Dust particles	10 times more
Sulphur dioxide	5 times more
Carbon dioxide	10 times more
Carbon monoxide	25 times more
Radiation	
Total on horizontal surface	15–10% less
Ultra-violet, winter	30% less
Ultra-violet, summer	5% less
Cloudiness	
Clouds	5–10% more
Fog, winter	100% more
Fog, summer	30% more
Precipitation	
Amounts	5–10% more
Temperature	
Annual mean	1–1.5° F more
Winter minima	2–3° F more
Relative humidity	
Annual mean	6% less
Winter	2% less
Summer	8% less
Wind speed	
Annual mean	20–30% less
Extreme gusts	10–20% less
Calms	5–20% more

Source: T. Walton Purdom (ed.). *Environmental Health*. New York, Academic Press, 1971, p. 219.

Temperature tends to be increased even though incoming solar radiation is reduced. Heat is re-radiated from buildings, streets and sidewalks, increasing the temperature, and the usual cooling effect produced by evaporation of water from plants and soil is greatly reduced. Increased temperature because of re-radiation is

also facilitated by the three-dimensional structure of a city, since solar energy is thus often reflected from one absorbing surface to another, maximizing conversion to heat. The decreased cooling effect due to lack of evaporation from plants and soils is further enhanced by rapid run-off of rain-water into sewage systems where evaporation is again lower than in the rural situation. Over-all, the cooling efficiency of the city is also reduced because of decreased wind speeds so that heat is less efficiently carried away.

The city also converts a great deal of energy from other forms to heat: domestic cooking, heating of buildings, air-conditioning systems, internal combustion engines.

Relative humidity in urban areas averages somewhat lower than in rural ones. The principal factors involved are the lowered evaporating rates and the higher temperature in cities.

Clouds, fog and precipitation, etc. are generally increased in cities even though relative humidity is decreased. This results from the release of the large amounts of particulates into the air over the city. Sources of particulates include cooking fires, dirt road transportation, industry and vehicle exhausts. These particulates serve as nuclei on which water droplets condense. Also, the warmer air over the cities sometimes results in upward movements of clouds into colder air, resulting in precipitation in areas outside the city.

Induction of rainfall

A great many cloud droplets must come together to form a single raindrop. In nature, this occurs through the presence of numerous ice particles (below —10° C) or water droplets (above 0° C) which serve as nuclei on which additional water condenses, forming the raindrops. Efforts to induce rainfall have utilized this principle. Techniques include dispersing (or seeding) solid carbon dioxide (dry ice) into clouds (at 0° C to —10° C), utilizing silver iodide crystals (at temperature below —5° C) or using large water drops (with cloud temperature above 0° C). These processes often induce precipitation, but the methods are primarily effective in increasing rainfall (by 15 to 20 per cent) from clouds in which precipitation is already occurring. In tropical and subtropical areas, silver iodide can be seeded into a cumulus cloud with resultant rapid development of the cloud. The effect is to produce fairly heavy precipitation over a small area. Particulate pollutants from cities sometimes also increase rainfall, but the over-all percentage of increase is small.

In summary, current potential for inducing rainfall by these methods is limited to slight increases (10 to 20 per cent) where precipitation is already occurring or conditions for precipitation are just right.

Fog dispersal and hail suppression

Fogs in which the temperature is below 0° C can be dispersed by using crystals of silver iodide, propane gas, dry ice or brine. Freezing then occurs, resulting in snowfall and clearing of the fog. The practice is fairly effective and is used at some major

airports. Unfortunately, a comparably effective technique has not been found for dispersal of warm fogs. Seeding large drops of water into such fogs has not been found to be consistently effective.

When cloud temperature is below freezing, ice particles may grow to considerable size and fall to the ground as hail. Thus, very large numbers of particles are sometimes introduced to serve as nuclei for precipitation in the hope that hail formation will be prevented. Most efforts have had limited success at best, but some success has been achieved by shooting silver iodide into clouds with rockets and grenades. There is also some evidence that supersonic sounds suppress hail and that shock waves from explosions may shatter hailstones. Regardless of effectiveness, rockets are fired into thunder clouds for this purpose in numerous countries, including southern Africa, U.S.S.R., United States and Kenya.

Air pollution control[1]

The first step of a pollution-control programme is to measure the concentrations of pollutants in the air. Most often, oxides of sulphur and solid particulate matter (dust, smoke, etc.) are measured. These are usually chosen because they are relatively easy to measure. Other pollutants often measured are oxides of nitrogen, organic compounds and carbon monoxide. In Singapore, monitoring is done by the Anti-pollution Unit. It is able to monitor sulphur dioxide, dust and smoke, carbon monoxide and oxides of nitrogen. Singapore has twenty-two monitoring stations to measure sulphur dioxide and smoke and fourteen stations to measure total dust fallout. The thirty-six stations are distributed throughout industrial, commercial and residential districts.

Numerous monitoring stations carefully distributed lead to the second major step of a pollution control programme, i.e. identification of major sources of air pollution (industries and power plants using coal or oil, motor vehicles).

Monitoring and source identification provide the basic data for air pollution control.

In Singapore, the government passed legislation on air pollution control in the form of the Clean Air Act (1971) which applies to industrial and trade premises. In addition to the Act, the government passed certain regulations in 1972 outlining the limits on the discharge of air pollutants.

Briefly, the following points should be considered if air pollution control is to be undertaken effectively in a developing country: establish the means to monitor air pollution; determine major sources of air pollution; devise standards of air quality; determine technological and economic feasibility of standards; evaluate appropriateness of air quality standards utilizing environmental, health and economic criteria; provide for government regulations and enforcement of air quality standards; provide continuous monitoring of the air to ensure that increases in pollutants are identified and located at once.

1. See Chapter 5 on 'atmosphere' in Part II, page 117.

Additional alternatives that should be pursued for improved air quality include: use of fuels producing lower quantities or less toxic emissions; collection of rubbish and centralized burning of it with use of air pollution control devices; maintenance of private automobiles in proper operating condition with minimum emission of smoke and fumes; use of alternatives to the internal combustion engine for automobiles; increased use of public mass transport and decreased use of private automobiles.

Water treatment and purification[1]

Besides the various treatments of sewage and those dealing with water purification, the mass algal culture process is an alternative treatment process which may be used for recovering nutrients from wastes as high protein livestock feed and for generating water fit for irrigation. It offers the following benefits: water pollution is checked in the urban areas; urban water resources are conserved by reclamation and re-use of water; polluted water is converted into a more acceptable form for use in irrigation; an animal feed is provided without the use of conventional water resources; an additional animal food resource is provided, utilizing nitrogen otherwise wasted; protein supply is increased without making additional use of the limited agricultural land resources; livestock feed is produced continuously in spite of seasonal and climate changes.

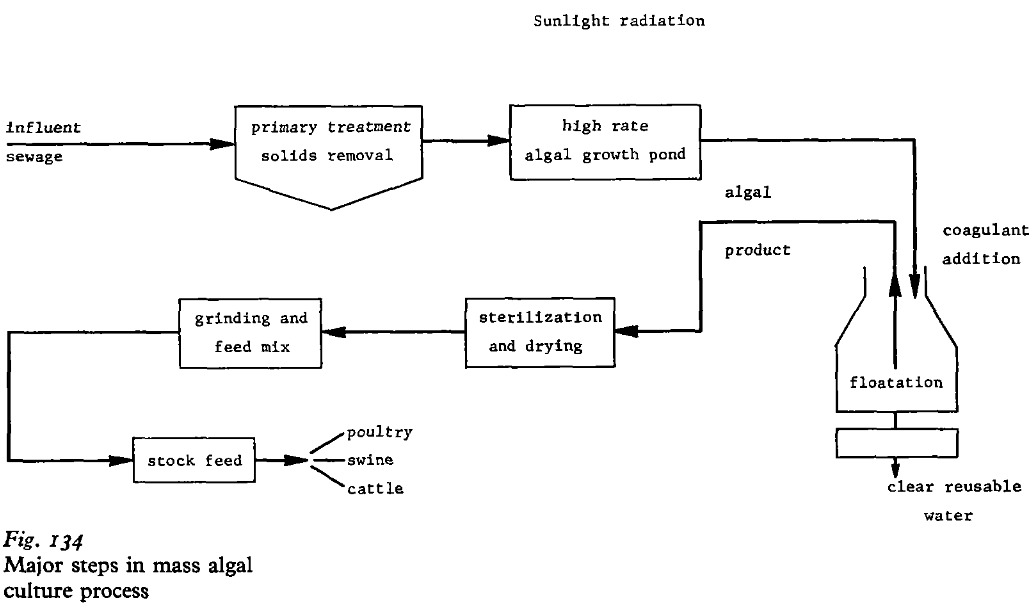


Fig. 134
Major steps in mass algal culture process

Fig. 134 shows the major steps of the mass algal culture process. Influent sewage may be either domestic or industrial in origin. First, the solids in the sewage are

1. See Chapter 6 on 'water' in Part II, page 139.

removed by sedimentation. The relatively clear fluid which remains is pumped into a high rate algal growth pond on a continuous basis. It is important that this pond be shallow to ensure penetration of sunlight to the bottom of the pond to aid in the growth of the algae. The alga grown is a unicellular organism of the genus *Chlorella*, and it produces annual dry yield of 112,000 kg/ha (100,000 lb/acre). It has high protein content and rapid growth.

The algae are removed from the liquid medium by adding a coagulant, which causes the algal cells to form flocs. These flocs can be readily removed from the liquid by floatation. The flocs are then dried, and the dried remains are ground and mixed with other animal feed for direct feeding to livestock.

When the algal growth in the liquid medium is harvested, various components including magnesium, calcium, nitrate, ammonia, phosphorus, carbon and oxygen are also removed. Thus the reclaimed water with reduced salt concentrations appears to be suitable for irrigation, but not for human consumption. Table 90 gives high rate pond water characteristics after an algal harvest. In general, irrigation water having total dissolved salts (TDS) below 500 mg/litre may be used on all but the most sensitive crops, while TDS concentrations of 500 to 1,500 mg/litre may be used widely, provided there is adequate drainage and leaching. However, studies indicate that the TDS value of 264 mg/litre appears to be satisfactory for use on most soils.

It is important to note that the reclaimed water must be disinfected before use in irrigation, or the feed grown on land irrigated by the effluent water must be cooked before being eaten. The disinfectant commonly used is chlorine gas or dry powdered hypochlorite.

Table 90
High rate pond water characteristics after algal harvest

Dissolved solids (in mg/litre)			
HCO_3^-	232	Ca^{++}	31
Cl^-	58	Mg^{++}	11
SO_4^{--}	88	Na^+	120
PO_4^{---}	0	K^+	30
Organic nitrogen	3.4	NH_4^+	51
Fe	0	Cr^{++}	0
TDS value[1]	381.4		243 = 624.4 mg/l
		Volatile	425
Suspended	8	Settleable	0

1. Total dissolved salts.

Source: Proceedings of the Ninth Session of the Regional Conference on Water Resources Development in Asia and the Far East. United Nations, 1971.

Control and management of waste

In the civilized world of today man is the only animal which creates waste at an exorbitant rate. An estimated total of more than 4,300 million tons of solid waste was produced in 1969 in the United States.

Waste, mainly a problem of the developed countries, is now reflected glaringly in dumps, in rivers, in coastal waters and on beaches in developing countries. Accumulations of such materials are becoming a great problem both as aesthetic disasters and health hazards. When burnt they contribute to air pollution; when thrown in water they raise the level of BOD; water percolating through them pollutes groundwater supply; and when left alone they are excellent breeding grounds for disease bearing organisms such as rats, cockroaches and flies.

Sources and kinds of waste

The main types may include: organic waste covering general sewage, animal and plant residues, food and fibre processing materials and other oxygen demanding residues; inorganic waste comprising salts, minerals, mineral acids, silt and sediments, fine suspending metals and their compounds and other chemicals discharging into streams or on the soil from various metallurgical and chemical industries, mining operations, and other natural resource development activities; in this category plant nutrients and sediments washing into streams are posing major problems; radioactive waste results from mining and refining of radioactive minerals such as uranium and thorium, atmospheric fall-out from testing nuclear weapons and accidental or unauthorized release of radioactive isotopes from power reactors or from research laboratories; thermal waste results from numerous industrial processes and electric power generators using water for cooling.

As to its origin, waste can be classified as follows: waste of immediate biological origin such as faeces and decomposed or decomposing bodies of animals and plants; waste of agricultural farms and processes; wastes of industrial processes including mines and fossil fuels; waste of domestic and commercial origin.

Thermal waste

It is fast emerging as a serious source of water pollution. The situation is now being aggravated by nuclear or atomic reactor plants because of their lower thermal efficiency as compared to most efficient conventional plants burning oil or gas. Also, there is almost no heat loss to the atmosphere from the atomic plants. Thus, at present stage of technology, there is more heat discharge to the cooling water per kilowatthour of electricity generated by the nuclear plant. Normally there is a temperature difference of 6°–8° C between inlet and outlet water for an electric power generating plant. When the resultant temperature of water is much above the surrounding air temperature, fogs or mists may develop.

Except for some bacteria and blue-green algae, little life survives in very hot waters. In general no higher organisms can live normally in waters above 35° C. Fish is, in fact, rare even above 30° C.

To control thermal pollution, two major devices include cooling towers and cooling ponds. With each of them, the cooling water is recirculated after it has been treated or allowed to stand for the heat to be dissipated. There is a water loss through evapor-

ation with consequent increase in salinity in cooling towers, and ultimately there is a problem of disposal of this water. Cooling ponds, on the other hand, require large areas for successful dissipation of heat and are thus feasible where land is cheap. However, the flow of exit water must be controlled in order to safeguard the usefulness of natural waters.

Radioactive waste

This waste differs from other industrial waste in that it cannot be detected by human senses. Its toxicity is greater and there is no known method for destroying it. It does present the most difficult problems for disposal until its natural decay or disintegration.

Three major management concepts in the control and disposal of this waste include: 'concentrate and contain', which are achieved by volume reduction and isolation in controlled areas such as underground storage tanks, salt mines and bottom of the sea away from man and his natural resources; 'dilute and disperse', in which case radioactivity is reduced to acceptable levels by dilution in air and water; 'delay and decay' approach can be used in the discharge of low-level liquid waste into the ground where it has no chance to contaminate the groundwater. Safe disposal is possible after standard treatments.

Solid waste and its management

Solid waste is considered here to include garbage (food wastes), rubbish (paper, plastic, wood, metals, tin cans, crockery, glass), demolition products (bricks, masonary, piping and lumber), sewage treatment residue (sludge and solids from the coarse screening of domestic sewage), dead animals and other discarded materials. *Per capita* daily waste production in a community depends upon the dietary habits, life styles and living standards of its members. The degree of industrialization and urbanization also influences waste generation. In urban India it ranges between 0.25–0.5 kg as compared to 0.6 kg in Switzerland, 0.82 kg in the United Kingdom and 2.5 kg in the United States. These figures do not include industrial waste. In India, total solid waste production in urban areas (population 20 per cent of the total 547.9 millions as per 1971 census) ranges between 11 and 20 million metric tons per year. This amount is estimated to increase at a rate of 1.33 per cent per year. In the United States the daily *per capita* waste production (2.5 kg/day) is expected to double by 1980.

Table 91 shows some physico-chemical values of typical municipal waste at source and at disposal site for Calcutta and at source for some other cities. Calcutta refuse contains large proportion of green coco-nut shells. (This may also be true for Sri Lanka and Malaysia.) Table 92 shows composition of typical municipal waste.

At present there is little attention devoted to the management of waste in most developing countries. At best disposal system involves collection of waste from city centres for open burning and uncontrolled dumping at many sites or flushing through water ways; each adding burden to our already polluted air, water and soil.

Conservation and rational use of natural resources must seek to avoid waste at all

Table 91
Physico-chemical values of municipal waste in some metropolitan areas

Item % by weight	Calcutta (average of 308 samples)		Poona	U.S.A. (1963, average of 4 cities)
	Source	Disposal site		
Garbage	45.14	47.25	67.64	22.0
Paper	3.18	0.14	8.74	42.0
Glass	0.379	0.24	0.58	6.0
Rags	3.6	0.28	1.63	0.6
Plastics	0.645	0.54	0.72	1.5
Moisture	—	—	—	21.0[a]
Carbon	19.28	20.31	22.03	28.0
Hydrogen	—	—	—	3.5[a]
Nitrogen	0.564	0.573	0.823	0.33
Sulphur	—	—	—	0.16[a]
Phosphorus	0.591	0.451	0.596	—
Potassium	0.427	0.479	0.544	—
Density (kg/m^3)	470.0	540.0	298.0	250.0
Calorific value (Hcv in kc/kg)	1 500.0	—	1 700.0	2 728.0

(a) M. A. Benarde. *Our Precarious Habitat.* New York, W. W. Norton & Co., 1970, p. 164.
Source: Refuse disposal studies at Calcutta. *Technical Digest* (Nagpur), No. 15, March 1971, p. 1.

Table 92
Composition of typical municipal waste

Contents	Percentage
Moisture	21
Cellulose, sugar and starch	46.6
Lipids (fats, oils, waxes)	4.5
Protein, 6.25N	2.0
Other organics (plastics, etc.)	1.15
Ash, metal, glass, etc.	24.93

Source: M. A. Benarde. *Our Precarious Habitat.* New York, W. W. Norton, Inc., 1970, p. 164.

cost. In this context waste implies destruction or loss without equivalent gain and occurs both in production and consumption. Steps, therefore, must be taken to prevent such waste at all stages of mining and industrial operations.

Some of the significant projects and processes for re-entry of wastes into resource cycle and at the same time reducing the bulk of unsalvageable waste for ultimate economic disposal include:

Pig feeding and fish rearing. In Malaysia and other South-East Asian countries, Chinese farmers collect garbage from markets, restaurants and hotels, etc. under contract and use it as pig and chicken-feed for pork and poultry production. The waste accrued from piggeries and poultry is in turn channelled into the fish pond for fish production.

The system seems to operate profitably, but research and studies are however needed to put it on more economic and sanitary footing.

Organic waste. Various processes and experiments are already under way in various parts of the world in order to manage municipal waste and its organic materials. They include: chemical conversion of wastes into saleable materials and how to chemically reduce the volume of solid waste in order to simplify their collection and disposal; pyrolysis of solid municipal waste, i.e. thermo-chemical conversion of complex organic solids into combustible gases, tarry liquids and a carbon residue. The process reduces the volume of waste and at the same time produces marketable by-products, e.g. carbon residue and combustible gases for fuel and tarry liquids for use as pesticides as well as a source of organic acids; possibilities to develop human food and animal feeds from waste accruing from the food-processing industries (protein concentrates).

Salvage industry. It has been in operation in several countries for many years: rags, paper, glass, tin cans and other scrap metal, etc. are sorted out for recycling. The process is profitable where labour cost is no problem and the scrap industry has been supplementing needs, especially in copper and steel production. In India, the Central Leather Research Institute (Madras) has developed a process to manufacture leather boards from waste leather.

Composting. Potentially it is a very desirable method for re-using waste and conserving our natural resources. Metals, glass and similar inorganics are sorted from mixed refuse and the remainder is converted to sanitary organic fertilizer and soil conditioner under controlled microbial degradation. Although composting is not new, it is only recently that major improvements in the process have been effected to allow large-scale disposal. With increasing demand for food and with decreasing soil fertility, the use of compost to improve agricultural production becomes critically important. Compost is used for crops such as sugar beets, potatoes, tomatoes, cucumbers and citrus. When used with low organic soils it can improve the water holding capacity and further help decrease the leaching of chemical fertilizers through the soil.

Sanitary landfill or tipping. Sanitary landfill is defined as 'a method of disposing of refuse on land without creating nuisance or hazard to public health or safety, by utilizing the principles of engineering to confine the refuse to the smallest practical areas; to reduce it to the smallest practical volume, and to cover it with a layer of earth at the conclusion of each day's operation or at such more frequent intervals as may be necessary'. In contrast, the open dump is a great nuisance and a health hazard, it being infested with disease vectors. Three methods are normally employed for sanitary landfills: area, trench and ramp methods. With the area method, refuse is deposited in horizontal layers on relatively flat ground, compacted and covered over sides and top with soil or another inert material. With the trench method land is excavated in the shape of a trench in which refuse is deposited and compacted and then covered with earth. Separation of refuse by walls of soil (Fig. 135) checks fires from spreading

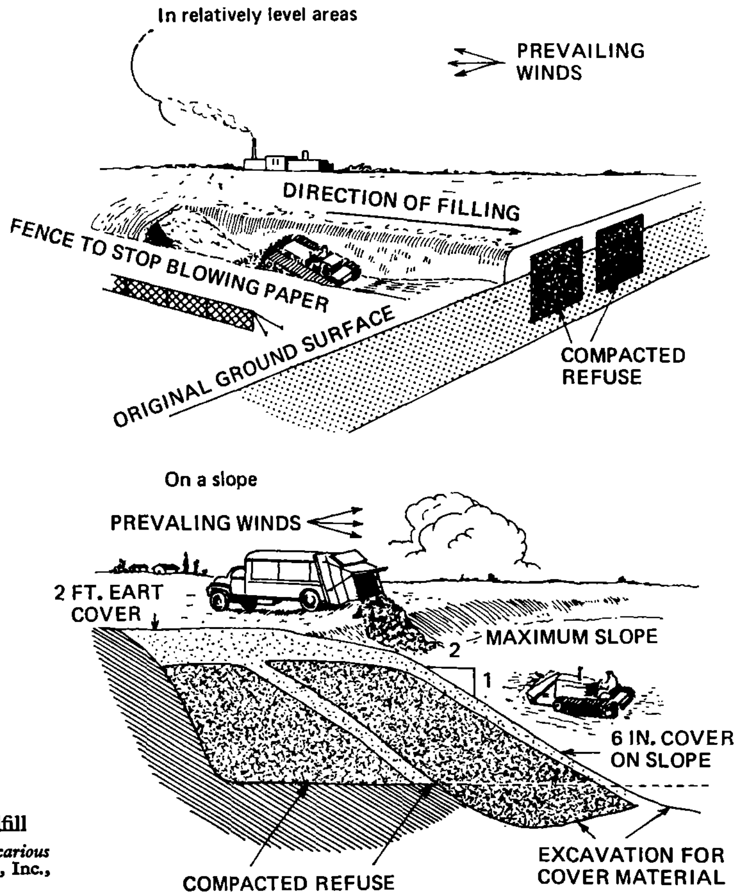


Fig. 135
Schematic representation of a well-planned sanitary landfill
Source: M. A. Benarde. *Our Precarious Habitat*. New York, W. W. Norton, Inc., 1970, p. 158.

beyond a single cell. The ramp method, on the other hand, employs existing ravines or quarries in which refuse is deposited on an angle against the side of the ravine or quarry. When refuse is buried, chemical changes set in as a result of microbial activity. This raises the internal temperature as high as 150°–160° F well above the tolerance range of pathogens, thus assuring that the process is not a health hazard. Instead it helps in the reclamation of land. However, leaching and subsequent seepage or run-off water soluble chemicals may be a problem and hence deserves increased technical and scientific study for elucidation.

Incineration. This method of disposal seems particularly suitable for urban areas where waste burying sites are rather scarce. The method involves reduction to ashes of combustible material by controlled burning at high temperature and hence the most efficient, hygienic means of disposing of all types of combustible waste. As a result it can be located closer to communities than landfills. Other advantages of incineration are reduction in the volume of the refuse to an economic minimum and elimination of putrescibles. The residue is the clean ash which can be used for landfills. The heat

generated during the process can also be utilized for various operations including steam and electricity production.

Summary

There are three basic options in dealing with waste materials: dump them untreated into the nearest convenient environment such as the air, a river, a lake, a well, soils or ocean; contain and treat them within a delimited environmental waste management park where engineered semi-natural ecosystems such as oxidation ponds, spray irrigated forests and landfills do most of the work of decomposition and recycling; treat them in artificial chemo-mechanical regeneration systems. The first operation is based on 'solution by dilution' and has been the main system for disposal in developing countries. With increased intensity in community's consumption and discard, this option must be abandoned as fast as possible.

The second option proves more economical when large areas are available. Such open spaces will not only help preserve our environmental quality but also provide other uses, e.g. feed and fibre production. For self-contained waste management, 500 to 5,000 ha of land may be needed. Re-cycled water and recovery of useful products from wastes may more than compensate for the cost of the land.

The third option involving abiotic treatment and re-cycling is especially suited to densely populated industrial areas. A significant problem of managing for quality in the aquatic environment is that of developing suitable methods of sewage treatment which will remove nitrates, phosphates and other nutrients efficiently and cheaply.

19 Regulation of fertility

Introduction

As regards fertility, a human couple may be faced with one of two problems. One that they have or are having more children than they want; the other, is that they are unable to produce any children or as many children as they may want. The problems of infertility in males and in females are discussed in Chapter 16, page 357. The present chapter is primarily concerned with the biological problems of birth limitation.

Throughout recorded history human beings have controlled births through such measures as contraception, abortion, infanticide, castration, sterilization, regulation of sexual behaviour before and during marriage, the encouragement of celibacy and homosexuality, and magic. To be sure, the earliest methods of birth control lacked the performance qualities which intelligent life would be expected to demand. Evidence suggests that human societies have struggled to achieve high quality birth control measures ever since population pressures were felt, and the process continues today.

The main criteria for birth control are effectiveness, acceptability and harmlessness. The destruction of newborn children was an effective form of birth control, but it was certainly not harmless, and as human cultures evolved, infanticide became unacceptable to modern civilizations.

Today there are numerous alternative methods of birth control from which individuals may choose or which societies may impose. The effectiveness and harmlessness of birth control methods can be quantitatively assessed. Acceptability is a qualitative standard and involves such variables as morality, comfort and convenience. One social group may impose long periods of abstinence and the rhythm method on moral grounds, while another society may condone unrestricted sexual freedom in association with universal contraception, also on moral grounds. The criteria of effectiveness, harmlessness and acceptability may in the judgements of both social systems be satisfactorily achieved.

Contraception

Contraceptive methods of birth control include all those procedures which allow a fertile couple to engage in sexual intercourse without allowing impregnation to occur. Four general categories of contraceptives, in decreasing order of efficiency, will be discussed. In the first category, the classical or combined oral pill stands alone. The second category includes the intra-uterine device (IUD) and the sequential oral pill. The third category includes such devices as the condom, cervical cap and diaphragm with chemical spermicides. The fourth category comprises such unreliable contraceptive procedures as rhythm, jellies and foams, vaginal douching and coitus interruptus. At the end of this section the effectiveness of contraceptive methods will be compared.

The combined pill

The breakthrough in oral pill contraception occurred in the 1950s when Gregory Pincus observed that a progesterone-like drug when administered to female rabbits prevented conception. The research was extended to humans with unusual success in terms of effectiveness, acceptability and harmlessness. Over the years the 'pill' has been improved to the point where its effectiveness is now essentially 100 per cent; its acceptability in terms of cost, comfort and convenience continues to increase, and with dosage adjustment and comprehensive studies of medical side effects, its safe use by the vast majority of women in the world is assured. Nevertheless, there are some risks to the good health of women who take the pill. But these risks appear to be no greater than the hazards to health associated with pregnancy or cigarette smoking.

The combined pill is a mixture of artificial ovarian hormones which resemble, in chemical structure and physiological action, the natural hormones, progesterone and estrogen. These natural and artificial hormones are made up of steroid compounds. Fig. 136 illustrates the carbon numbering convention for steroid molecules. Estrogens contain eighteen carbon atoms, androgens (male hormones), nineteen and progesterone, twenty. A natural estrogen (17 β-estradiol) produced by the human ovary appears in this figure with two synthetic estrogens which are derived from plant sources and used in combined pills. The structural formula of progesterone is shown (D) with one of the many laboratory-prepared progestins also used in combined pills. The yam plant provides the chemical raw material from which most progestins are derived. Typically a pill contains less than five milligrammes of progestin and a small fraction of estrogens. Variations in proportions of progestin to estrogen, and variations in the specific structure of the synthetic hormones combined are mainly accommodations to women whose physiological reactions to hormonal pills vary.

A woman takes the pill orally every day beginning five days after the onset of menstruation and continuing through the twentieth or twenty-first day. The effects of the pill can be described at a number of levels. However, the two main effects are: they prevent ovulation and they regularize the menstrual cycle even for those women who normally do not experience regular cycles.

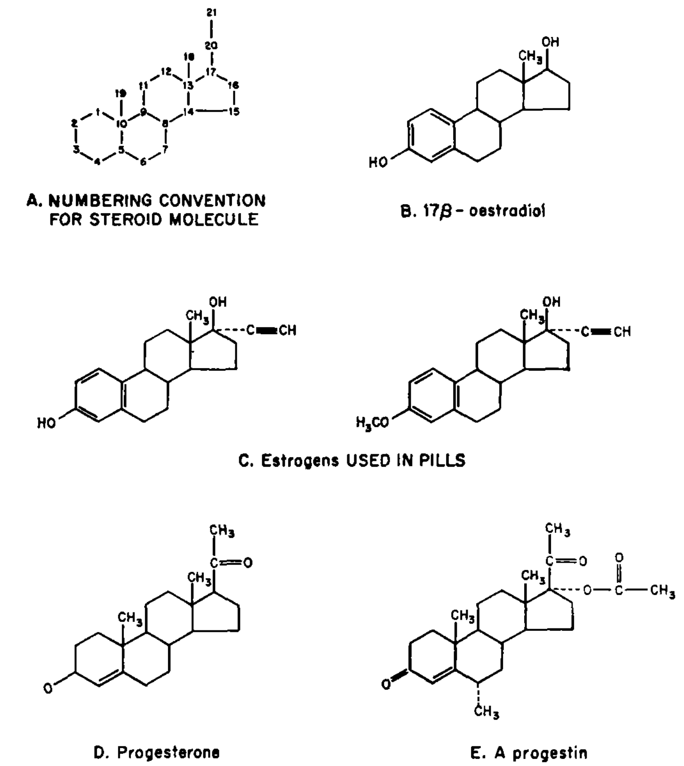


Fig. 136
Structural formulas of natural and synthetic steroid hormones

The progestins and estrogens in combination prevent the secretion of the gonadotrophic hormones by suppressing the hypothalamic releasing factors. The inhibition of the production of FSH and LH prevents follicular development and the shedding of eggs. In the early days of the pill progestins alone were administered. This often resulted in irregular bleeding, a condition corrected by the cycle-regularizing estrogen. In addition to this action on the pituitary-ovarian system through the hypothalamus, the estrogen-progestin combination alters the nature of the cervical mucus and renders the cervix hostile to penetration by sperm, changes the rate of tubular transport of the egg, and modifies the character of the endometrium making implantation unlikely. Thus the pill provides multiple efficiency as a contraceptive. To all intents and purposes the pill is 100 per cent effective, with most failures having been attributed to women who have forgotten to take tablets regularly in the course of the cycle. The risk of pregnancy rises with the number of pills missed.

The clinical side effects of the pills have been the subject of intensive medical research ever since their development. Among the useful side effects of the pills are: a reduction in the incidence and intensity of the pains, nausea and faintness associated with menstrual bleeding (dysmenorrhea); an alleviation of premenstrual tension, irritability and anxiety; a reduction in blood and tissue loss during the period of menstrual bleeding; relief from excessive body hair growth caused by male hormone expression;

relief from acne (pimples) caused by inflamed sebaceous glands; relief of pain of ovulation (mittelschmerz) and increase in sexual appetite and feeling of well-being. Catalogued with the annoying effects of the pill are: a usual gain in weight due to a possible disruption of glucose metabolism; nausea during the first several cycles; an enlargement and tenderness of the breasts, and a possible increase in the number of headaches.

Serious questions are being asked and investigated about the long-term effects of oral contraceptives. The fears that the compounds of the pill are carcinogenic have thus far not been supported by the findings. There may, however, be a relationship between the pill and blood-clotting or thrombosis. Studies in the United Kingdom in the late 1960s reported that women taking oral contraceptives were more likely to die from the effects of inflammation of deep veins together with blood clots, or pulmonary embolisms than women who did not take the pill. Against the effect of the combined pills, drug laboratories produced a variety of pills with reduced quantities of estrogen, the component which appeared responsible for thromboembolic disorders. The relationship between oral contraceptives and thromboembolic disease is not definitely established. Studies in the early 1970s now suggest that there are no significant differences in the incidence of thromboembolic diseases between women who took the pill and women who did not regularly use the oral contraceptive.

Contrary to some popular belief, there is no evidence that the reproductive life span of women who take the pill will increase. In the absence of developing follicles and corpora lutea, the ovaries become smaller and there is no storage of eggs. Women appear to experience the menopause at the normal time. Upon withdrawal from the pill, the ovary begins its normal functioning, although menstrual rhythm requires several months to return. Almost all women ovulate by the third post-treatment cycle.

The sequential pill

The sequential preparation differs from the combined pill in that for the first fourteen to sixteen days the composition of the tablets is estrogen alone, and for the next five days is a mixture of estrogen and progestin. To avoid confusion the woman takes inert pills (placebos) for the eight days she does not take the hormone tablets. The sequential pill imitates the endocrine background of the normal menstrual cycle and some women respond more physiologically to a hormonal sequence than to the steady double dose of the combined pill. The sequentials produce less changes in the tissues of the uterus and the cervical mucus than do the combined pills which imitate pregnancy. The milder action of the sequentials accounts probably for their slightly higher failure rates than the combined pills.

The intra-uterine devices (IUD)

Knowledge of the contraceptive effects of placing foreign objects in uteri possibly antedates the ancient practice of Arab and Turkish camel drivers who placed stones in the uteri of camels to prevent pregnancies during long journeys. The use of IUD in

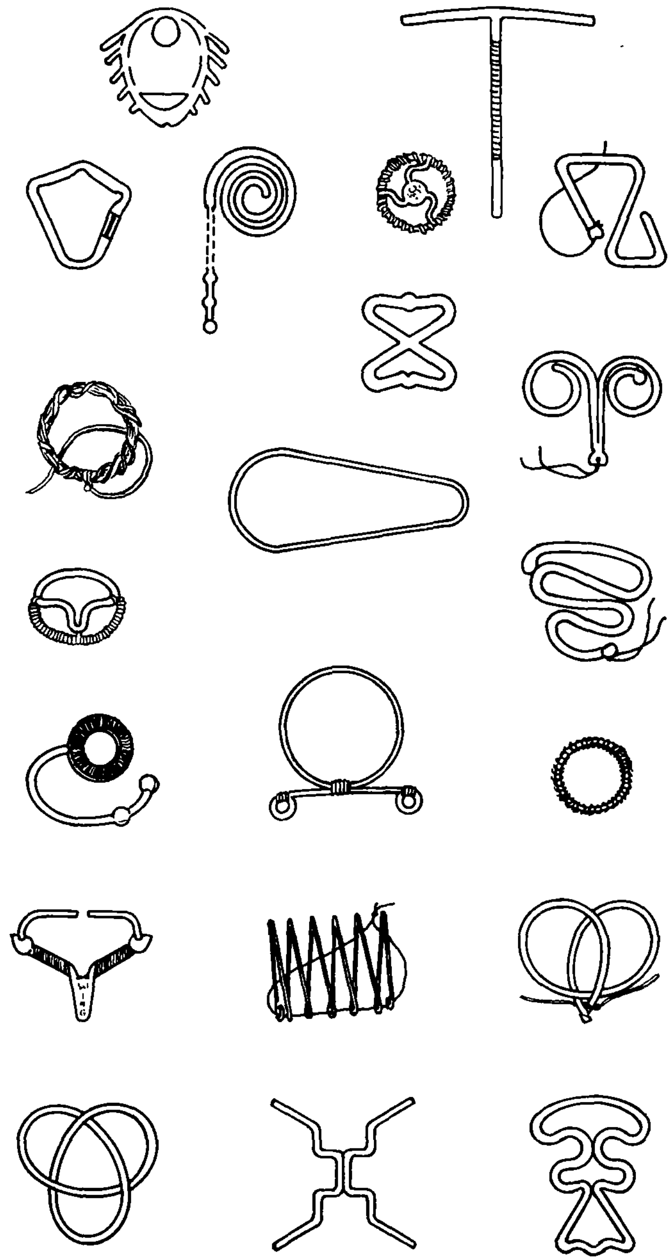


Fig. 137
Types of intra-uterine devices

humans has a long history, but the extensive and successful use of the IUD as an effective, acceptable and harmless contraceptive is recent.

The IUD of today is a relatively simple polyethylene plastic or stainless steel configuration which fits into the cavity of the uterus. Types of IUD are illustrated in Fig. 137. The great numbers of different designs reflect a trial and error search for efficiency. The earliest problem with IUD was to create designs which would prevent

inadvertent expulsion of the device—usually during menstruation. To enable the woman to be assured that the device was in place, and to assist the physician in its removal, threads which hang outside of the vagina were attached to the IUD. Some IUD without threads contain impregnated magnets which through the use of a magnetometer, reassure their presence. Since some IUD are also opaque to X-rays, their positions and presence can be confirmed. When it was discovered that the contraceptive effect of the device is proportional to the contact surface area of the IUD, efforts were directed toward designs that increase surface contact. Recently, the presence of copper in the uterus was found to have excellent contraceptive effects as an enzyme inhibitor, and research with copper impregnated IUD is proceeding.

Various sizes and designs of shields, as in other types of IUD, are available as appropriate for the woman. Women who have not borne children have smaller uteri than mothers and require smaller IUD.

Several explanations of the mode of action of the IUD are current. Most probably the IUD changes the uterine environment making it hostile to the sperm cells and the cleaving egg. Contact between the IUD and the inner wall of the uterus stimulates an increase in the number of white blood cells. The more contact, the more white blood cells. Since white blood cells are known to phagocytose a significant number of sperms in the normal uterus, the IUD stimulated uterus may be thoroughly gameto-

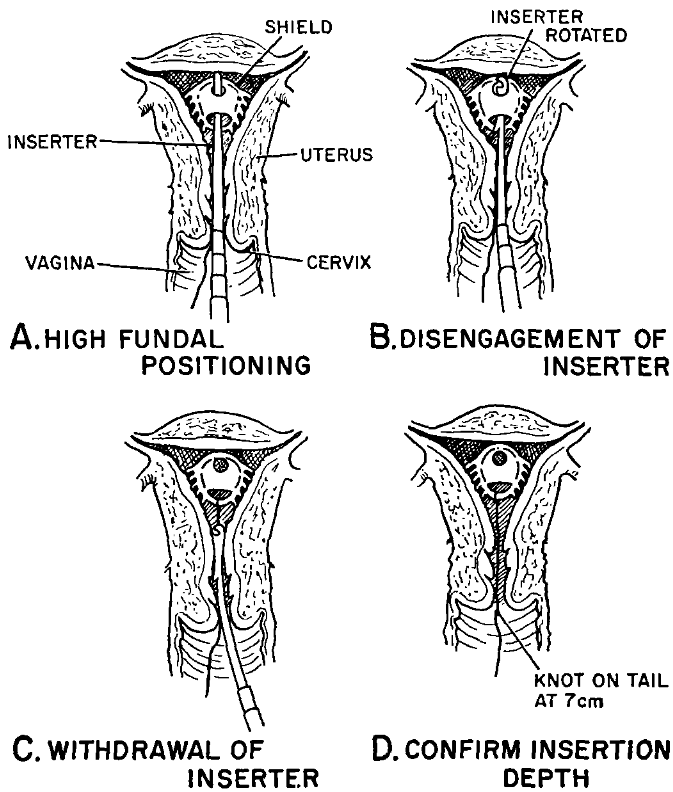


Fig. 138
Insertion of Dalkon shield IUD
Source: Davis. *Intra-uterine Devices for Contraception.* New York, Williams and Wilkins, 1971.

toxic. A second explanation of IUD action, based on infra-human studies, suggests that IUD stimulates the rapid passage of the egg through the oviducts. Since the fertilized egg arrives in the uterus before the endometrium is prepared it cannot implant.

The greatest dangers of IUD are associated with insertion. To be certain that the recipient is not pregnant, IUD are inserted during or within several days after menstruation. Although insertion is a relatively simple procedure, it requires the skill of a physician or trained paramedical. Perforation of the uterus by the instrument carrying the IUD sometimes occurs. The insertion procedure for the Dalkon shield is illustrated in Fig. 138.

The most common side effect associated with women with IUD in place is bleeding between menstrual periods and cramps. However, these effects usually persist for only two or three months, for those women who can endure this period of adjustment. Where pregnancy has occurred in women with IUD, they have gone to term with the IUD remaining outside of the embryonic membranes. Rates of spontaneous abortion in these situations have been slightly higher than in women with no IUD in place.

The IUD is an inexpensive reversible contraceptive favoured by women who do not wish to submit to a regime of pills, or who for medical reasons cannot take hormonal pills.

The condom

Throughout the world, the condom is the most widely used mechanical method of contraception. It is a rubber, plastic or animal tissue tube closed at one end. The condom fits tightly over the erect penis during intercourse and retains semen after ejaculation. When properly used, it is an effective contraceptive and prophylactic against venereal disease.

The major objections to the condom are that it interferes with the pleasures of sexual intercourse by insulating the glands against direct contact stimuli, and forces the interruption of foreplay while the condom is rolled on the erect penis. The recent development of silicone lubricated rubber condoms, and the thinner plastic (ethylene ethyl acrylate) condoms provide increased sensitivity during intercourse.

Hundreds of millions of condoms are manufactured annually in automated factories. In the modern process of manufacture each condom is tested for holes and weaknesses before it is sealed. No longer is it necessary, or even advisable, to test a condom by insufflation before use. However, rubber deteriorates in time, and the shelf life of a rubber condom is about five years. The new plastic condoms have an indefinite storage life.

Condoms are easily distributed since they require neither prescriptions nor fittings by medical personnel. Their contraceptive use is indicated when: the female cannot use contraceptives; the male, for psychological reasons, must control reproduction; the female suffers from certain vaginal diseases; the male, disturbed by premature ejaculation, wishes to delay orgasm, and the couple engages in sporadic or infrequent

intercourse, as when partners are frequently separated. No medical side effects are known to result from the use of condoms.

The diaphragm

Throughout history people have invented wide assortments of devices that fit into the vagina of the female and act as barriers to sperm. The diaphragm, a rubber cup with a rubberclad rim of flexible steel, is the most recent and probably the last of such devices. Ranging in diameter from 45 mm to 105 mm it is inserted into the vagina to fit over the cervix. The diaphragm must be fitted by a physician and the woman must be instructed thoroughly on the skills of use. To be at all effective, diaphragms must be used in conjunction with spermicidal jellies or creams. Before insertion these chemicals are coated on the edges and on the concave underside of the diaphragm. In effect, the diaphragm is merely a vehicle to hold the chemical spermicides.

In use, the diaphragm is inserted prior to intercourse, and removed between six and twenty-four hours after intercourse. It does not interfere with the pleasures of coitus nor are there any health side effects for those who find the diaphragm acceptable.

The cervical cap

The cervical cap is a thimble-shaped rubber or metal barrier to sperm, 22 mm to 31 mm in diameter, which fits closely over the cervix. It is maintained in position except during menstruation when the woman removes it. For effectiveness it should be used in conjunction with chemical spermicides. Apart from the advantage over the diaphragm of removing the necessity for physical preparation before intercourse, placement of the cervical cap requires considerable skill on the part of the female. Its popularity has declined rapidly in recent decades.

Chemical contraceptives

The modern chemical contraceptive consists of a spermicidal substance and a carrier that also tends to act as a barrier to the passage of sperm into the vagina of the female. Jellies and creams used without such mechanical barriers as condoms, diaphragms and cervical caps, are not effective. Foam tablets, when placed in the vagina, react to produce a barrier froth of CO_2 which distributes the spermicide throughout the vagina. An aerosol with a butane propellent also has the effect of creating a relatively impermeable wall of spermicidal substance. Chemical contraceptives when used alone have not yet demonstrated failure rates low enough to recommend their extensive use. However, their use is far better than no contraceptive at all.

The rhythm period

The rhythm or 'safe period' method of birth control restricts sexual intercourse to periods of physiological sterility in each menstrual cycle. As defined by the Roman

Catholic Church, it is not an unnatural contraceptive method. Rhythm is the only form of birth control—save abstinence—of which this religious group approves for its 600 million members. The method depends upon the fact that the one egg shed during each cycle is fertilizable for a period of less than twenty-four hours, and that sperm can remain alive in the female reproductive tract for approximately forty-eight hours. There is then a mid-cycle period of about seventy-two hours which is not safe for intercourse if pregnancy is to be avoided. For safety's sake couples are advised to add several days of abstinence on both ends of the seventy-two-hour period.

The main problem with the rhythm method arises in identifying the unsafe period. Menstrual cycles are by no means the same or regular for all women. They may vary from twenty days to forty days with ovulation dates ranging from the seventh to the twenty-seventh day of the cycle. For a couple successfully to avoid intercourse during the fertile period, the woman must keep an accurate calendar record of her cycle on an annual basis. The problem of irregularity of the cycle is further complicated by the degree of uncertainty attached to the precise day of the cycle during which ovulation occurs. A practical method of determining the time of ovulation is based on the phenomenon that progesterone increases the metabolic rate with a consequent rise in body temperature (about 0.6° F). The woman charts her daily morning temperature (before rising) using a rectal thermometer. Intermittent fevers will confuse the data. Noting the pattern of ovulation over a period of many cycles, the couple can predict their safe and unsafe periods. Needless to say, the rhythm method requires intelligence, organization and high motivation. For couples with such characteristics and with women who are predictably regular, the method works well. The rhythm method does not work for sexually active menopausal women with irregular cycles.

Coitus interruptus

The withdrawal of the penis from the vagina just before the occurrence of ejaculation is probably the oldest contraceptive practice known. This method has been criticized as being unnatural, unfulfilling and ineffective. However, recent studies have demonstrated low failure rates for coitus interruptus among couples of many societies who have been able to adjust to the procedure. Failures have been attributed to preorgasmic secretion of sperm, but statistically the effectiveness of coitus interruptus compares favourably with mechanical contraceptive methods. The psychological problem of adjustment to coitus interruptus are sufficiently compelling to counsel against it in preference to less strenuous methods.

Measures of effectiveness

Contraceptive methods should be evaluated on their use-effectiveness rather than on calculation of their theoretical or laboratory effectiveness. Use-effectiveness, then, is the capacity of a contraceptive to prevent accidental pregnancies, and is expressed as the

failure rate of the contraceptive per one hundred women years of exposure (HWY). The formula for calculating this failure is:

$$\text{Failure rate per HWY} = \frac{\text{Total accidental pregnancies} \times 1{,}200}{\text{Total months of exposure}}.$$

(Use-effectiveness can be determined by other methods that are more reliable but much more complex than the application of this formula.) The denominator describes the length of time the woman was exposed to pregnancy. Months of nonexposure due to pregnancy, illness or separation are deducted. In the numerator, the number of unplanned pregnancies are prorated to 100 years of exposure. A failure rate of ten can mean, then, that for a given contraceptive in use by one hundred women for one year of exposure, or for four women for twenty-five years of exposure (or one woman for one hundred years of exposure) ten unplanned pregnancies occurred. For an individual woman choosing to use a contraceptive with a failure rate of eight throughout her lifetime of twenty-five years of fecundity, she may expect 2.0 accidental pregnancies. The pregnancy rate per HWY assuming no contraceptive practice is eighty. This can be interpreted to mean that among one hundred women using no contraception, eighty can be expected to be pregnant by the end of one year.

It must be emphasized that failure rate figures are statistical, and describe populations, not individuals. Further, failure rates for a given contraceptive may vary considerably with the socio-economic, educational and motivational levels of the group under study. Failure rates make no distinction between patient failure and

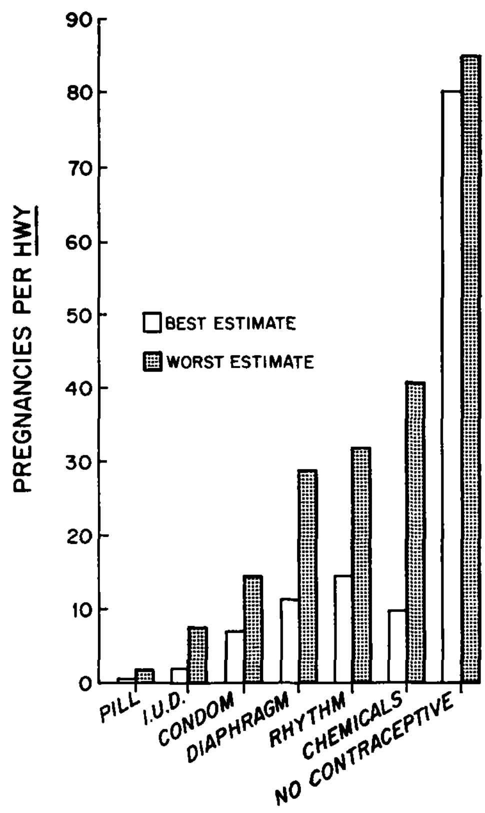


Fig. 139
Pregnancy rates of contraceptive methods compared with rates associated with non-use of contraceptives
Source: Peel and Potts. *Textbook of Contraceptive Practice*. Cambridge, 1969.

method-failure—any unplanned pregnancy is considered to be a failure regardless cause. Research procedures for the determination of failure rates ideally involve clinical trials rather than retrospective demographic studies which depend largely upon recollection for data. The research literature is replete with often widely conflicting failure rate claims for contraceptive methods. Fig. 139 roughly summarizes the failure rate range of some selected contraceptive methods based on studies which were comparable in methodology. The reader must be cautious to note, for example, that of the approximately twenty-five different kinds of modern IUD, reported measures of failure rates range from 0.4 to 18.3. Fig. 140 presents a simple over-all comparison of the use-effectiveness of the major contraceptives and no contraceptive use. Points on each curve predict the probabilities of pregnancy by months of exposure to the risks of pregnancy. For example, after ten months of unprotected sexual intercourse, 78 out of 100 women have become pregnant, and about nine out of 100 women whose husbands were condom users (or so they claimed) became pregnant.

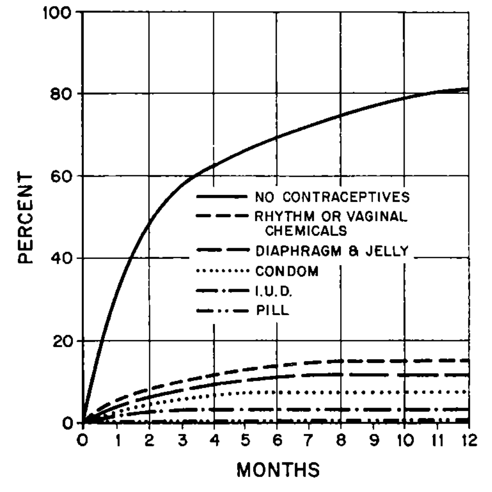


Fig. 140
Estimated percentage of women becoming pregnant by months of active sexual intercourse and by contraceptive used (Courtesy Philip Jones, M.D.)

Future prospects for contraception

The development of contraceptives has been a continuous process. Currently, there are contraceptives about to become generally available, others are in early clinical testing stages, and still others are under laboratory development. Most of the new contraceptives are intended for use by women, and many are variations of the pill.

Most promising are contraceptives which provide low continuous doses of progestins. They can be administered orally daily as 'mini-pills', injected intramuscularly every three to four months, or contained in a silicone rubber time capsule which is embedded beneath the skin and which releases progestin at constant rate. The small dose of progestin is believed to have the effects of altering the consistency of cervical mucus and thus preventing the entrance of sperm to the uterus, affecting the lining of the uterus in such ways as to prevent implantation, and/or cause ova to

travel too quickly through the oviducts. Small continuous doses of progestins do not necessarily prevent ovulation. The advantages of these new contraceptives include their high effectiveness (a reported failure rate of the 'mini-pill' is 3.7 HWY), their reversibility, the ease and acceptability of their administration, and their relatively mild side effects. The major side effect is irregular vaginal bleeding which is treated by the administration of estrogens.

The so-called 'morning-after pill' is a procedure to prevent the implantation of a possibly fertilized ovum soon after an unprotected exposure to pregnancy. The woman takes large oral doses of estrogens for three to five days. Their probable effect is to upset the critical balance of pituitary and ovarian hormones which exist after ovulation and disrupt normal transport and implantation. This treatment is usually discomforting to women and is not to be recommended on a regular basis. However the search continues for a safe pill, with a similar effect, which might be taken as a routine, once a month.

In other areas of biochemical research, substances are being sought with contraceptive properties without the drawbacks of the artificial steroid hormones. Moreover, long lasting spermicidal substances which can be plugged into oviducts, or contained in vaginal rings are under development.

The knowledge that some forms of pathological infertility have an immunological basis, suggests the possible utility of immunizing a woman against her husband's sperm, or the male against his own sperm.

Little progress has occurred in the development of new male contraceptives. Steroids and other chemicals which have effectively provided temporary sterility have had the disadvantages of reducing libido or producing dangerous side effects. Clips that close the vas deferens and inserted plugs which occlude the vas deferens, offer some hope as alternatives to sterilization, but thus far their failure rates are high.

Sterilization

There are several reasons for contemplating sterilization for an individual. The person may carry genes for physical disorders which he or she would not want to pass to offspring. Also it can happen that a female, for medical reasons, should not risk the strain of pregnancy and child birth. In recent years, sterilization as a means of birth control has become popular throughout the world and many relatively simple and effective medical procedures have been developed to provide sterilization for both males and females.

For the female, the most common operations are to cut and tie off the Fallopian tubes (tubal ligation), remove portions of the tubes, or most effectively, remove both tubes in their entirety (bilateral salpingectomy). The effect of these operations is to prevent the egg from passing from the ovaries to the uterus, and prevent the passage of sperm to the egg. The best time for the operation is soon after delivery when the uterus is enlarged and the Fallopian tubes are close to the abdominal wall. The operation involves a small single incision along the midline of the abdominal wall.

Other techniques enable surgeons to perform tubal ligations by the vaginal route. In addition to cutting and tying tubes there are procedures for occluding the tubes physically and chemically. A recently developed technique involves the use of the laparoscope, a long slender tube which is inserted through a small incision in the navel. A cold light passes down the tube enabling the operator to look through the tube and observe the reproductive organs. For sterilization a second incision is made below the navel and a forceps carrying an electric current is inserted and manipulated in view of the laparoscope. The physician cauterizes the Fallopian tubes with the forceps. The operation requires twenty minutes and the woman suffers minor discomfort for several hours.

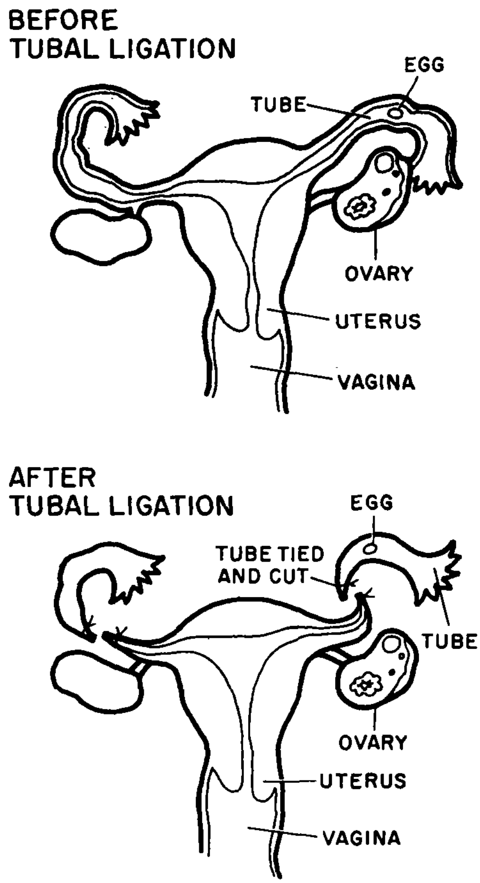


Fig. 141
Tubal ligation

There are no known side effects associated with contraceptive sterilization involving the impairment of the passage of eggs through the Fallopian tubes. Since neither the ovaries nor the uterus are removed normal female ovarian and menstrual cycles continue unchanged.

The sterilization procedure for the male involves the cutting of the vas deferens (vasectomy) at locations close to the surface of the skin, usually at a point where

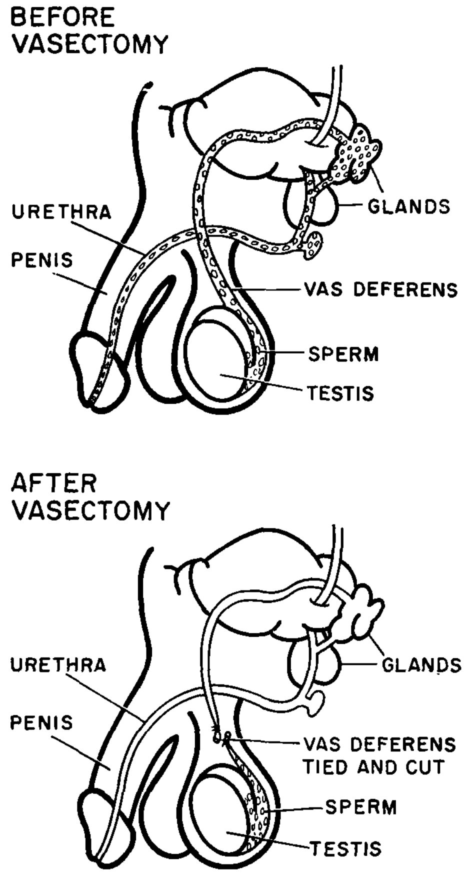


Fig. 142
Vasectomy

the scrotum nears the body. Using a local anesthetic, the surgeon finds the cord by feeling it and lifts and holds it in a fold of skin. Making an incision over the cord he applies forceps to the cord about 2 cm apart, cutting away the intervening portion. Longer pieces (6 to 8 cm) of the vas deferens may be excised to ensure more effective sterilization while reducing chances of reconnecting the cut ends should the patient wish to reestablish his fertility. The exposed ends of the cords are tied with silk thread and sometimes cauterized with an electric needle. The procedure is then repeated on the other side of the scrotum. Variations of the procedure occur.

Diagrams of the operative procedures for tubal ligation and vasectomy appear in Figs. 141 and 142 respectively.

As in the female, sterilization through the severence and occlusion of the ducts which transport sex cells does not affect male physiology. The testes continue to function normally producing sperm and hormones, and sexual performance continues with the single exception that sperm cells will not occur in the ejaculated semen. However, spermatozoa may exist in great numbers 'up stream' from the excision and conventional contraceptive methods should be practised until masturbatory semen

samples fail to reveal any motile sperm. Thirty per cent of vasectomized males show viable sperm eight weeks after the operation and in one instance living sperm were observed in semen one year after a successful vasectomy. Those sperm cells which continue to be produced in the seminiferous tubules of the testes and remain in the ducts up to the sealed vas deferens are phagocytosed by macrophages, the large ameboid phagocytic blood cells in connective tissue.

Chief among the side effects of vasectomy is a brief dragging discomfort in the testes due to the loss of support to testes from the vas deferens which formed part of the spermatic cord. In approximately 5 per cent of the cases a swelling, usually temporary, due to extravasated blood (an hematoma), forms.

Pregnancies after sterilization have been attributed to recanalization of Fallopian tubes or vas deferens, residual spermatozoa or the overlooked presence of a third vas deferens.

Sterilization should be regarded as a permanent and irreversible operative procedure, with individuals arriving at the decision only after careful consideration of their future plans. The repair through rejoining of the Fallopian tubes is a long and delicate operation and the chances of success are not encouraging. Males who wish to ensure against a change of mind, can consider storing frozen samples of their sperm in a sperm bank. It may even be possible for vasectomized males to have sperm removed from their testes to be used, as with frozen sperm, through artificial insemination.

Psychologically, for most persons studied, sterilization increases their pleasure in sexual relationships by removing fears of pregnancy. In general, sterilization is performed only after the birth of a number of unwanted children. This phenomenon was true of 95 per cent of those interviewed in a study in Switzerland.

Abortion and its medical implications

Defined as the termination of pregnancy before the twenty-eighth week, abortion is classified as either induced or spontaneous. Induced abortion is the most widely used method of controlling fertility throughout the world. In some countries one out of two conceptions is terminated by induced abortion, and in several countries, the number of induced abortions has actually exceeded the number of births. High rates of induced abortions even occur in societies where effective contraceptives are readily available.

Spontaneous or natural abortions account for the termination of between 10 and 20 per cent of all pregnancies counted after the first missed menstrual period. If spontaneous abortions were to include losses prior to the first missed period, evidence suggests that the frequency would be greatly increased. Chromosomal abnormalities account for about 30 per cent of these natural terminations of pregnancy. For humans, and other mammals, spontaneous abortion is an important regulatory mechanism which protects the species against overwhelming numbers of defective infants.

The medically approved methods of induced abortion are dependent upon the stage of pregnancy. Pregnancies of less than twelve weeks may be terminated by a

procedure known as dilation and curettage, or with an instrument which removes the foetus and its tissues from the uterus by aspiration. In the dilation and curettage operation the cervical opening is enlarged usually by metal dilators. Using curettes the physician scrapes the lining of the uterus removing the foetus and its tissues. The aspirator, after dilation, accomplishes the task of curettage more rapidly and with less bleeding. Each method is statistically safer to the woman than full term pregnancy. The medical procedures for terminating pregnancies between twelve weeks and twenty-eight weeks usually involve the cutting open of the uterus (hysterectomy) or the injection of hypertonic solutions of salt or glucose into the amniotic cavity of the foetus. After twelve weeks of pregnancy, the risks of complication and death from abortion increase considerably. After twenty-eight weeks, in the eyes of civil law in most parts of the world, the foetus is viable and capable of surviving outside of the mother, and destruction of the foetus is illegal.

A large percentage of the induced abortions performed throughout the world are illegal and unsupervised by medical authorities. The techniques of nonmedical abortionists include high pressure vaginal douches of soapy water or antiseptic, the piercing of the foetal membranes with hat pins or other sharp instruments, or the administration of such drugs as quinine or ergot in doses intended to be lethal to the foetus but sublethal to the pregnant woman. Needless to say, illegal abortions are often fatal. Among the reasons for the high incidence of illegal abortion are legal and moral judgments that abortion is a kind of murder that should never be permitted, or be permitted under specified circumstances short of the free choice of the woman who seeks the abortion. Some argue that even contraceptive practice, by interrupting the continuity of life, is murder, while others may argue that at any age a foetus is only a potential human being whose rights and privileges begin at birth and whose destruction prior to birth is the free choice of the woman who bears it. Many current laws declare the foetus to be a human being from the time when the mother begins to feel its movements. And while lawmakers, clergymen and physicians argue, desperate women with unwanted pregnancies seek relief from clandestine practitioners. Modern trends are for considering abortion acceptable under certain circumstances to avoid the birth of unwanted children and some countries have reviewed their laws to make such abortions legal.

As a birth control measure, induced abortion is inferior to contraception. In fact, the incidence of abortion is a measure of contraceptive failure. As contraceptive practice becomes universal and achieves essentially 100 per cent effectiveness, the need for abortion should practically disappear.

Prostaglandins

Until recently no chemical had been found that acts as a genuine abortion-inducing drug (abortifacient). One of the newly discovered classes of hormone-like substances, prostaglandin $F_2\alpha$, appears to be capable of inducing abortion at any stage of pregnancy. The drug seems to cause the regression of the corpus luteum and the subsequent termination of pregnancy. Research studies are now investigating the action of the

drug when administered intravenously, orally and vaginally. It may some day be possible, for example, for a woman to insert into her vagina a tampon treated with a prostaglandin after a missed period, and abort the implanted embryo.

Socio-economic and cultural aspects of birth control[1]

Any policy which aims to introduce or to extend family planning and (or) birth control measures in a population must take into account several considerations which are peculiar to the socio-economic and to the cultural levels of this population.

Economic considerations

In many rural communities of developing countries, children either help their parents in their daily routine or add to their families income by working and earning wages. Whenever this is the case, children are not only objects of preference for their own sakes, but they are also desired as economic assets. Child labour is not illegal in some parts of Asia. In addition many undertakings are family businesses in most parts of these countries. To own a business in this manner one has to make sure there is enough manpower and having more children is the best way to be sure of the supply.

Sometimes special circumstances lead people to want as many children as they can. The following two examples taken from a paper submitted to the second Asian population conference, held in Tokyo in 1972, may make the point clear.

In one of the Asian countries, a vital proportion of the rural community lives within an increasingly complex economy of rubber small-holdings.

These smallholders are now living within a complex of government rubber and crop-diversification development activities. Many are felling old rubber trees, replanting with high yield ones, with government grant or assistance, and trying to maintain their livelihoods by crop diversification during the five to seven years before the new trees become matured enough to be tappable. Each element in the material (financial and otherwise) support they receive from government for replanting, will crucially govern how they will examine the proposition that they have only two or three children. The basic new perspective in their thinking, 'in seven years . . .' may quite conceivably lead them—very intelligently in their own interests—to decide that this is most exactly the time to produce as many children as they can with the hope that those who survive will be able as soon as possible to help tap the new trees.

In several Asian countries, increasingly large numbers of rural families are now entering the frightening as well as uplifting experience of Agrarian Reform—frightening because many of the paradoxical forms of security they knew under bondage to the landlord are no longer available, and the real support of government agrarian reform and other services like new

1. For all relevant data about family planning and birth control (first attendances and contraceptive methods accepted, age-distribution and average age of acceptors, acceptors by education and by economic level, over-all results of birth-control programmes on population growth rate, etc.), the reader may refer to periodic national and regional population census, United Nations surveys and to the publications of the United Nations and of the Population Council (N.Y.).

cooperatives, etc., are as yet unproven... It is easy to be aware that they need many children so that those surviving can help them in their newly owned (or leaseheld) fields, just as quickly as possible.

In these two examples and similar other cases, asking and expecting people not to have more than one or two children in view of the 'population explosion' is like asking them to abandon their own pursuit of happiness and security. However, economic considerations do not always lead people to desire many children. When people see that large numbers of children cannot be of help in gaining and keeping a comfortable standard of living, and when, for any reason, parents level of aspiration for their children rises, they begin to think about limiting the number of their offspring.

Authority considerations

Several societies in developing countries (e.g. in Asia) are characterized by male-orientated patriarchal cultures. It places power and authority in the hands of the senior male members of the household, kinship group and the society at large. Thus, to keep the power and authority in one's own kin group and to continue one's own name, most people strongly desire male children. To enjoy more of the privileges that come by having sons, some people beget as many children as they can, and sometimes people who have female children keep on reproducing in the hope of finally having a son.

In almost every traditional African or Asian society, a woman's most triumphant day is the one on which she gives birth to a boy.

Survival considerations

In many parts of Africa, Asia and Latin America, due to the lack of proper and adequate public health and sanitary services, many children die at an early age. In the year 1970, out of every 1,000 children born in Turkey, 155 died before reaching one year of age. This figure was 142 for Pakistan, 139 for India and 127 for the Khmer Republic. They indicate a very high infant mortality in the first year of life, but when the mortality of children up to an age of 6 or 7 is also considered, one can appreciate the parents' anxiety in these countries and their efforts to reproduce as much as they can. They want to make sure that even after death takes away some of their children, they will still have some left.

Welfare considerations

In many developing countries, where old age pensions and economic security are not yet established, the care of elderly parents by their young and active children is expected through their culturally inculcated values and sentiments. Thus often an expressed motive, for wanting more than one or two children, is to secure aid and comfort during one's own old age.

When a social system has devised ways of curbing people's anxiety about their

financial security during old age, be it through pension, national insurance or any other security, then consideration may be given to the welfare of the children. Parents, who are reasonably sure that their children can have a good chance to survive and that in their own old age they will not need to depend on their children, begin to be concerned with their children's education and comfort, during and after their period of economic dependence on them.

Cultural attitudes and educational perspectives on birth control

Sometimes sentiments and preferences do not correspond to changes in social and economic conditions; be these familial, local, national or global. Nevertheless such inappropriate sentiments and preferences exist and persist because they are learned and people cling to them by habit, not because they serve any adaptive function. World population is increasing but social institutions and people's preferences and practices have not adjusted themselves to the demands of this new situation. If the change in sentiment and behaviour is to keep abreast of changing conditions, then people should be helped to learn and acquire new preferences and adopt new courses of action.

'Population education' is the term coined for the conscious efforts to provide this help. This effort has been described '. . . an educational programme which provides for a study of the population situation in the family, community, nation and world, with the purpose of developing in the students rational and responsible attitudes and behaviour towards that situation.' The content of any educational programme should have two sub-components of attitude and knowledge.

The attitude sub-component aims at changing an old attitude for a desired new one. Thus an ultimate intention of population education would be to create in the audience or the target group an understanding of the relationship of population dynamics to quality of life in its individual, family community, national and global dimensions. This understanding will hopefully create an attitude favourable to the wise planning of the family size. Since population problems and the factors creating or sustaining them are not the same in all parts of the continent, educational planners should first find out and then pay particular attention to the distinctive problem of the country and the locality concerned (e.g. interrelation of factors, namely age structure, migration, maternal and child mortality and over-all community health conditions and services, etc.).

The knowledge sub-component aims at providing facts. If people, contrary to their wishes, still beget many children, then one cannot escape the conclusion that, most probably these people lack the right knowledge for birth control and efficient family planning. For example, people of Rapa (French Polynesia), who have more children than desired, seek to limit their numbers through the lead provided them by their cognition of the fertile period in women. This is based on a theory of reproduction quite different from that currently held by the medical science. Accordingly, they believe that a woman is most likely to conceive during the three or four days immediately following menstruation.

In such cases where knowledge is either lacking or the popular cognition is faulty,

then correct mechanisms of human reproduction and birth control should be taught to the interested public.

For the school age group, topics related to population problems can be included in the school curriculum, either as a separate unit or as illustrative or drill material in already established subjects such as social studies, biology, ecology, mathematics demography and other relevant subjects. For changing the attitude of adults in regard to population problems and family size, educational programmes can use the group discussion technique. This is to help small groups of people to identify, examine and try to come up with solutions to their common problems by way of face-to-face interaction. This technique has proved very effective in changing adult attitudes and can be used both with the general public and in the in-service training or short term educational programmes for policy makers and programme planners.

But along with the above techniques, the wide utilization of the media of mass communication should not be ignored by educational planners. Through mass media one may not hope to change the public's attitude but it is often an effective measure to spread the correct information.

20 Does man have a future?

Danger of extinction

There is an abundance of literature on the likelihood of a rather rapid extinction of *Homo sapiens*. There are some of these publications which appeared several centuries ago, but more which are recent, in the last three years prior to this statement. Among the causes of the death of the species, war, famine and disease are suggested amongst others. Totally committed thermonuclear war would not kill immediately the world's populations, but would be likely to destroy the more advanced cultures that entered the conflict. Subsequent fall-out of radioactive material would take its toll, but it is unlikely that this would entirely eliminate the species. This is not to minimize the effects, which would be terrible to consider, but to view the future of the species. Famine, by its nature, cannot destroy the species. However few the survivors, there would be some, and they would have food for existence, if not sufficient to flourish. Disease is similar in that the concentration of the people is a factor in the transmission of a communicable disease. As the concentration of people diminished with the spread of the disease, the effect of the disease would diminish. Some non-communicable diseases, such as those of pollution by heavy metals (mercury, lead), are lethal in many populations.

Is there any factor, then, which might bring about the extinction of the species *Homo sapiens*? With the increase in the rate of technological advance, there is cause for concern, though panic is not likely to facilitate problem solving. This concept is not limited to accidental, but spectacular, war. It is also inherent in the rapidly increasing use of materials with little knowledge of their immediate effects and no knowledge of the more subtle ecological ramifications. A comparatively sudden change in our environment is more likely to result in our demise than would any other factor, if we use the history of the extinction of other species as our guide.

It has been said that technological advance is now out of control. That is, even sure knowledge of an impending disaster would not likely lead to our averting it. Although this is perhaps too strong a statement, it does appear that some of our technological institutions have become so ponderous as to be difficult to manage. It si but a short extension to admit the possibility that a subtle change could be

introduced through the liberation of existing materials into the oceans, resulting in a shift in, say, the species of plankton which form the basis of the food webs of the oceans. Atmospheric oxygen concentrations are very much dependent on these organisms. In a similar manner, if we were to suddenly discover that a breakdown product of a widespread product was toxic to nitrifying bacteria, it is entirely possible that we could not remedy the problem before the ecological changes got out of hand. The above are not examples of expected changes, but are suggested to encourage thoughtful application of technology, which, in itself, reduces the risks involved. Note, too, that the examples are fanciful, since a reporting of fact in this case would be an exercise in futility.

In this sort of problem, time is important and a re-inspection of rates can be instructive.

In Pliocene times, pre-men had at least a million years to adapt to the spreading of the savannahs. With the beginning of the Pleistocene epoch came comparatively rapid environmental changes, the spread and recession of glaciers in a matter of 250,000 years and, in the later cases, as few as 50,000 years. In the recent epoch, only a 10,000 year span has passed and only relatively small changes in man are visible in that period of time. Physiological adaptation to the cold in Eskimos and Aleuts may not be due to genetic changes, in spite of the fact that the climate is very harsh in their environment. If that change is a result of genetic alteration, then it is a small change and one of degree, not kind. Since the rate of mutation is relatively small (about one or two changes per 100,000 replicates per locus) and most mutations are detrimental and recurrent, a long span of time is clearly needed for adaptation to new environmental conditions.

A much more rapid expression of suitable phenotypes for a new environment results from sexual recombination if appropriate genetic material already exists. That is, no technology can as yet produce, on demand, a given useful gene. The alternative, then, is to have the correct genes on hand as they are needed. This is known as pre-adaptation and results, in part, from the effects of balanced polymorphism, which tends to retain genetic variability in a species. Other factors, such as the establishment of a home base in primitive man and the medical technology of modern man, also contribute to pre-adaptation.

Modern technology also has an influence on the over-all complexity of our ecosystem. Much of our applied technology aims at simplifying our surroundings, eliminating, for example, competitors for our food, such as those insects injurious to crops or livestock. Further simplification results from the elimination of those pests that are hazardous to our health or simply annoying. But the means of doing even this has ramifications beyond our intention. For example, DDT has increased crop yields considerably since its introduction in the 1940s. This was the intended and desirable result. But it also appears in our own foods, occasionally in unforeseen ways, such as the contamination of commercial fish and birds. There is also some indication that it is responsible for the declining nesting success in eagles and ospreys. The food organisms of these carnivorous birds concentrate the DDT by their own feeding habits. When the eagles and ospreys eat the contaminated food, the normal synthesis

of egg shell components is impaired. The elimination of these species, or any others, is an unintended form of simplification of our ecosystem. As that system becomes simpler, its stability and dependability are reduced. It is no small matter that a species or organism, once extinct, cannot be replaced.

It should also be noted that the ecosystem on which we depend entirely (and must continue to depend upon) has considerable resiliency and does absorb the adverse effects of our actions remarkably well. This is, however, a function of its present complexity. If we continue to modify our environment even with the most thoughtful technological change, we run the risk of simplifying it to the point where its resiliency will decrease dangerously.

In technological change, we are considering rates involving only decades or less. In human evolution, we talk in terms of at least thousands or tens of thousands of years for major changes. This inequity in time span leads to the conclusion that pre-adaptation to the impending (and impredictable) changes, is the most likely, if not the only, means of surviving modern technology. Our most valiant efforts to predict the course of ecological change on a short-term basis for a relatively simple ecosystem degenerate into forecasts. To do this on a world basis always requires assumptions of unrealistic conditions, although the practice may lead to a better understanding of the nature of the ecosystem in which we live.

Since the direction of change is unpredictable, the course of action in preparedness is to retain maximum variability in our genetic make-up. We know that the genes we would define as harmful in our present environment will not necessarily be so in the new one. The sickle-cell gene provides ample evidence of the differential fitness of the same gene in different environments. The present trend toward general miscigenation may enhance survival probability for future generations.

If man survives

If the human population continues increasing and reaches the predictable number of about 7,000 million people by 2025, as has been suggested elsewhere in this book, then the environmental change rate will be considerably enhanced, if we are to make inferences from present knowledge. This is due not only to the fact that more people will be tampering with our limited resources, but also to the fact that *per capita* use of resources will probably keep increasing. The resulting environmental impact is said to be decreasing the 'quality of life' for the members of the technologically advanced societies, even where the micro-environment of the people is improving. Life expectancy is high, infant mortality low, and other indices of personal wealth and well-being seem to show that man has never been better off.

'Quality of life' has been ill-defined in most of its uses, but we might consider it in terms of human evolution. In evolution, leaving surviving viable offspring is the criterion of success. Those factors which increase this likelihood are favoured and those which do not are disfavoured. Survival of the species is, however, the outcome and appears to be more important in evolution than the survival of any one population.

We might re-define our criterion of success as the survival of the species. It appears unlikely that any species (including man) can survive indefinitely, so it is probably necessary to further qualify the criterion of success to indicate that the species which survives for the longest possible time is the most successful. To keep a high quality of life we need a rich, varied culture. On the other hand, confronted with unpredictable environmental change, we, as a species that has just begun to play the game, can ill afford to loose any of our genetic or cultural variability, our most valuable asset. Man is unique in that he can also profoundly influence the course and rate of environmental change resulting from his activities; and he is unique in that he may care to protect his, as yet, unborn generations. It is the hope that he will apply his ingenuity to do both, thereby increasing the probability of his extended survival.

Further reading

ABELSON, P. H. Waste management and control. *Science*, 152 (3720), 1966.

BARDACH, J. E.; RYTHER, J. H.; MCLARNEY, W. O. *Aquaculture. The farming and husbandry of freshwater and marine organisms.* New York, Wiley, 1972.

BAUM, B.; PARKER, C. *Solid waste disposal.* Vol. 1 Ann Arbor Science Publishers, 1973, 397 p.

BERELSON, B. *Family planning programs: an international study.* New York, The Population Council, 1971, 304 p.

CONWAY, G.; ROMM, J. *Ecology and resource development in South-East Asia.* New York, The Ford Foundation, Office for South-East Asia, 1973, 82 p.

DASMANN, MILTON and FREEMAN. *Ecological principles for economic development.* London, Wiley, 1973.

FARVAR, Taghi; MILTON, John P. (ed.). *The careless technology: ecology and international development.* New York, Natural History Press, 1972.

GEERTZ, C. *Agricultural involution. The process of ecological change in Indonesia.* Berkeley, Los Angeles, University of California Press, 1968, 176 p.

HAAS, W. G. L.; RIJNBERG, Th. F. (ed.). *Proceedings of the international symposium on integrated surveys and environmental management in developing countries.* Enschede, ITC/Unesco Centre for integrated survey, 1973.

HOYLE, B. S. (ed.). *Spatial aspects of development.* London, Wiley, 1974.

OWEN, O. S. *Natural resource conservation: an ecological approach.* New York, The Macmillan Co., 1971.

PEEL, John; POTTS, Malcolm. *Textbook of contraceptive practice.* Cambridge University Press, 1969.

PEREIRA, H. C. *Land use and water resources in temperate and tropical climates.* Cambridge University Press, 1973.

POPULATION COUNCIL. *Studies in family planning.* A monthly review, since April 1970. New York.

——. *Reports on population/family planning.* A review without regular periodicity since March 1972. New York.

——. *A manual for surveys of fertility and family planning: knowledge, attitudes and practices.* New York, 1971, 340 p.

RURAL CHINA. Report by the School of Engineering Science. University of Edinburgh, 1972.

SASSON A. *Développement et Environnement (faits et perspectives dans les pays industrialisés et en voie de développement)*. The Hague, Paris, Mouton, 1974.

SEGAL, E. J.; ATKINSON, L. E. *Mechanism of contraceptive action of intra-uterine foreign bodies*. New York, The Population Council, 1973, 28 p.

TALBOT, Lee M.; TALBOT, Martha H. *Conservation in tropical South-East Asia*. Morges, IUCN, 1968, 550 p.

VAN DYNE, G. M. (ed.). *The ecosystem concept in natural resource management*. New York, Academic Press, 1969.

WATT, K. E. F. *Ecology and resource management, a quantitative approach*. New York, McGraw-Hill, 1968.

WHITNEY, J. B. R. Ecology and environmental control in China's development experience. Philadelphia. *Annals of the American Academy of Political and Social Science*, 31(1), March 1973, p. 95–109.